DOCUMENT RESUME

ED 379 290 TM 022 632

AUTHOR Datta, Lois-ellin
TITLE A Matter of Consensus. Working Paper.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 17 Aug 94
CONTRACT 43-3JAJ-3-01051
NOTE 146p.; "Prepared for OERI at the suggestion of Dr. Laurence Peters."
AVAILABLE FROM OERI, 555 New Jersey Ave., N.W., Washington, DC 20208.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Decision Making; Educational Practices; *Educational Research; *Evaluation Methods; *Information Dissemination; *Participation; Research Utilization; Synthesis
IDENTIFIERS *Consensus Models; *Office of Educational Research and Improvement; Panel Reviews; Strategy Choice

ABSTRACT

In response to a recommendation that the Office of Educational Research and Development (OERI) adopt consensus panels such as those used by the National Institutes of Health and the Food and Drug Administration, this paper tries to represent the range of consensus panel applications and to identify the major considerations for OERI application. A pyramid of consensus strategies is possible, with a broad base of many systematic decision-making strategies (such as group discussions or syntheses of reviewer comments) and a top point with a few model consensus panel approaches that require extensive participant agreement and rigorous examination of the evidence. It is noted that there is merit in the recommendation to use consensus development processes, but it is recommended that OERI try the consensus panel strategy of the National Academy of Science report only on a trial basis because effective features of consensus panels in health research probably cannot be transferred to educational research and practice. Strengths and weaknesses of consensus strategies in general and of specific strategies are discussed. Seven appendixes consider methodology and various types of consensus approach, and provide supplemental information. Four attachments provide additional details. (Contains 133 references in Appendix 3.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 379 290

TM

Working Paper

"A Matter of Consensus"

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

By

Dr. Lois-ellin Datta
Datta Analysis
68-1769 Laula Place,
P.O. Box 383768
Waikoloa, Hawaii 96738
Phone/Fax 808-883-8670
E-mail 74434.1122@compuserve.com

Office of the Assistant Secretary
Office of Educational Research and Improvement
U.S. Department of Education

August 17, 1994 Version

TM022632

BEST COPY AVAILABLE

Note from The Office of Educational Research and Improvement, OERI

This working paper, "A Matter of Consensus" by Dr. Lois-ellin Datta was prepared for Sharon P. Robinson, Assistant Secretary of the Office of Educational Research and Improvement, OERI at the suggestion of Dr. Laurence Peters. Individuals undertaking such projects are encouraged to express freely their professional judgment. This report, therefore, does not necessarily represent positions or policies of the U.S. Department of Education, and no official endorsement should be inferred. This working paper is intended to stimulate discussion and the exchange of insights on the government's role in facilitating consensus in R&D agenda setting and R&D dissemination.

We encourage you to request copies or share reactions and additional information with Dr. Susan Klein, OERI, 555 New Jersey Ave., N.W., Washington, DC 20208, Tel. 202-219-2038, E-mail: Sklein@inet.ed.gov or with the author, Dr. Lois-ellin Datta 68-1769 Lauia Place, P.O. Box 383768, Waikoloa, Hawaii 96738, Phone/Fax: 808-883-8670, E-mail 74434.1122@compuserve.com

Note from the author: Dr. Lois-ellin Datta:

This paper was prepared under Education Department Purchase Order 43-3JAJ-3-01051, dated September 23, 1993 for $2,450. In the work statement, the Department noted:

"As OERI prepares for its new Institutes and Office of Dissemination and Reform Assistance, its staff will need to have a full understanding of the range of consensus strategies that can be used to identify promising directions for R&D support, including the development of R&D agendas as well as its dissemination decisions about particularly effective R&D-based solutions or federal educational policies. In addition to learning about these consensus development strategies, OERI will need guidance on how to select topics that are most likely to benefit from consensus work as well as from scientific opportunities."

Two deliverables were required: a draft report to be completed by November 9, 1993 and a final "consumer report on consensus strategies to facilitate the support and dissemination of R&D" to be completed by January 31, 1994. The final report was to include:

- description and analysis of the key types of strategies
- time allocated to implementing them
- costs to agency sponsors
- key contacts and available materials
- strengths of the strategies for specified challenges such as agenda setting and dissemination
- insights on how to select and structure topics for consensus/synthesis treatment
- advice on using these strategies for R&D planning, and for recommending promising and exemplary products, programs and practices for further development and dissemination

This paper is the final deliverable, the consumer report, with the consumer taken to be OERI and ED leadership planning the new Institutes and the Office of Dissemination and Reform Assistance. Given the time available, the paper does not identify every possible relevant instance of consensus building approaches or consensus panels, nor does it discuss all the relevant literature. It rather tries to represent well the range of consensus panel applications and identify major considerations for OERI application. In addition to the text, I provided two seminars on consensus panels. This is the first of four papers on dissemination commissioned in September 1993 by OERI. The other three topics are (1) getting feedback on the quality of R&D innovations (Dr. Patricia Campbell); (2) the role of technical assistance centers in promoting education R&D solutions (Drs. Brenda Turnbull and Mary Leighton); and (3) wise consumer selection and use of education R&D solutions (Dr. Thomas E. Backer). The specific context of this paper is a recommendation by the National Research Council that OERI adopt NIH and FDA style consensus panels as one of OERI's dissemination strategies.

"Most syntheses done by an individual or small group ignore some variables and interactions that other equally 'expert' persons consider important....[In an] informally structured consensus-building process with no consistent criteria and little attention to underlying assumptions and paradigms...the review and revision of knowledge is usually limited and research findings tend to be accepted totally or rejected....[It will be] of limited value if the literature and [reviewing] experts represent only the initial synthesizer's perspective."
Ward, 1983, pp. 37-38.

"In order to move ahead [in education] in an orderly fashion, consensus---rational agreement by researchers and educators---is needed." Reid, 1992, p.3

A MATTER OF CONSENSUS

TABLE OF CONTENTS

Page

Notes on this Purchase Order from the Office of Educational Research and Improvement and the report author, Dr. Lois-ellin Datta i

Acknowledgments iv

Introduction: "A Matter of Consensus" in Context v

Ten Questions (and Answers) About Consensus Panels 1

Recommendations 20

Appendix I: Methodology 22

Appendix II: Agency, Organization, and Individual Contacts 24

Appendix III: References 27

Appendix IV: Consensus Building and Consensus Panels -- Applications for Research Agenda Setting 39

Appendix V: Consensus Building and Consensus Panels and Dissemination of Research Knowledge to Improve Practice 46

Appendix VI: But Is This for OERI? Analysis and Recommendations 81

Appendix VII: A Tale of Two Looks at Reading Research Syntheses 87

Attachments:

Information from the Office of Medical Applications Research, OMAR

Information from the Administration for Health Care Policy and Research, AHCPR

EPA Guidelines on Consultation and Consensus Processes and Negotiated Rulemaking

Briefing Paper on the American Psychological Society - R&D Agenda Setting on the Human Capital Initiative

ACKNOWLEDGMENTS

Many people have contributed generously to the preparation of this report. Some have sent splendidly useful papers. Some have graciously given long interviews on their experiences with consensus panels. Some have gone to much effort to provide names, telephone numbers and backgrounds of excellent sources of information. These are gratefully acknowledged in Appendix II. Special thanks are due to Dr. Richard Anderson, Director for the Center for the Study of Reading; Dr. Thomas Backer of Human Interaction Research Institute; Ms. Ellen Blasiotti of the National Institute on Disability and Rehabilitation Research: Dr. Gary Houser of the National Stroke Association, Dr. Tom Kelly of the Environmental Protection Agency; Dr. Alan Kraut of the American Psychological Society; Dr. Paul Shekelle of RAND; Dr. Ronald Sczyckowski of MAGI, and Dr. Charles Sherman of NIH/OMAR.

Special thanks go to Dr. Susan Klein, the Project Monitor. She responded abundantly to requests for information. She initiated searches that have proven invaluable. She provided constructive, thoughtful comments after her careful reading of draft materials. She kept the four paper-writers well-connected. She arranged, through the kind offices of Betty Demarest, for a meeting with other OERI experts on educational change and dissemination, including John Egermeier, Jean Narayanan, Wesley Smith, Carol Chelmer, Ellen Holland, Sue Gruskin, Lucille Reifman, and Eve Bither (among others whose input is also gratefully acknowledged). Thanks also to Peter Ellman, an OERI intern from the University of Pennsylvania, for providing useful insights into AHCPR and for helping with the camera ready preparation of this paper.

Last, but by no means least, Dr. Laurence Peters found the resources for the four papers examining new directions OERI might take in helping R&D make a good difference in education. He selected the topic of this paper and he provided encouragement and insightful questions throughout the effort.

INTRODUCTION "A Matter of Consensus" in Context

By Susan S. Klein, U.S. Department of Education

This report by Dr. Lois-ellin Datta was commissioned by the Office of Educational Research and Improvement (OERI) in the U.S. Department of Education (ED) to guide the implementation of its new legislation, the "Educational Research, Development and Dissemination Improvement Act of 1994" in Title IX of Goals 2000. Thus, opportunities for improving the federal role in education research and development (R&D) provide the primary context for this exploration of "A Matter of Consensus."

OERI has two immediate responsibilities that require substantial consensus to succeed. They are:

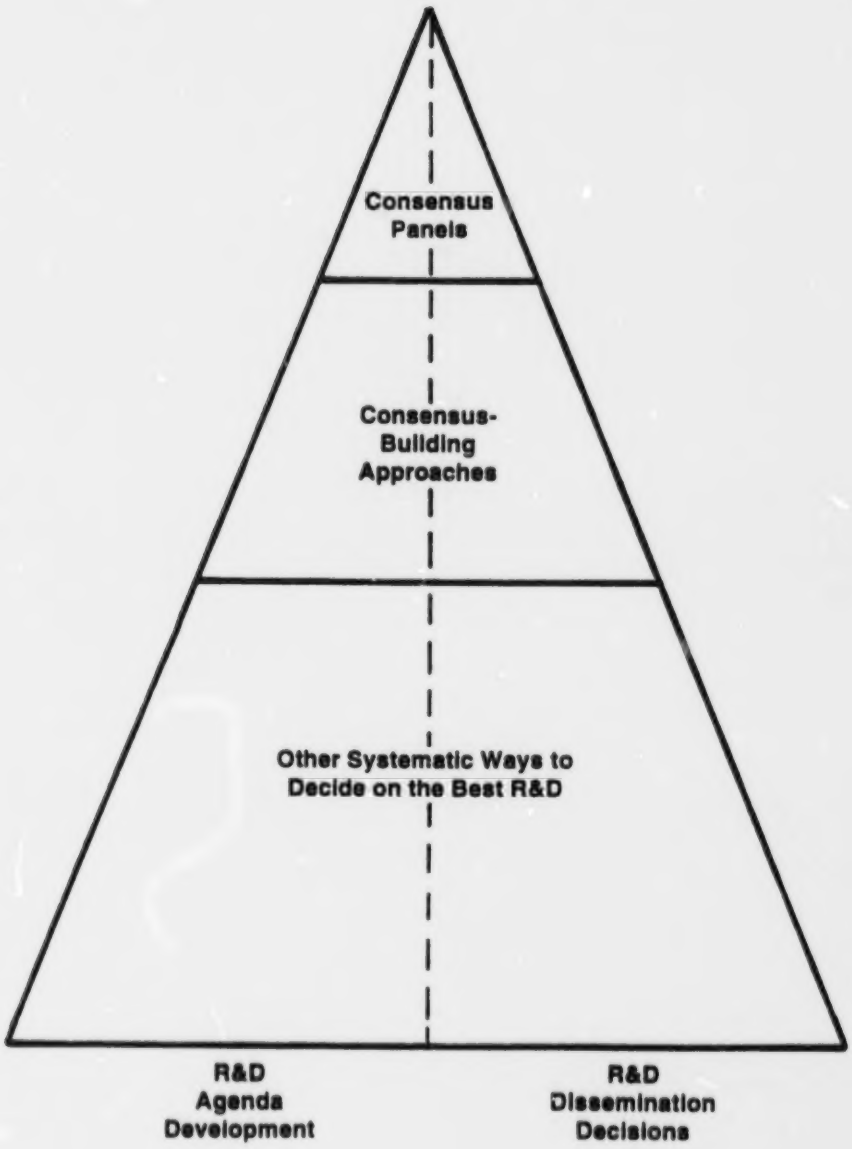
1. Planning national agendas to provide a sense of direction for research and development and to show how it may improve education and
2. Identifying and sharing the best R&D-based materials, programs and ideas that the nation has already produced so that education consumers can make wise choices of what will work best for them.

One type of consensus strategy has already been recommended by a distinguished Committee on the Federal Role in Education Research. In its 1992 National Academy of Sciences report, *Research and Educational Reform: Roles for the Office of Educational Research and Improvement* this Committee recommended that OERI implement a consensus development process to augment its more traditional reviews of the research and evaluation literature. It suggested funding consensus processes for two or three controversial topics a year at $400,000 to $900,000 each.

Following up on this recommendation, OERI asked Dr. Lois-ellin Datta, the author of *A Matter of Consensus*, to examine consensus approaches for the above two types of R&D decision making more extensively and intensively than the members of the National Academy's Committee were able to do. Since Dr. Datta, is a former high level official in OERI and in the U.S. General Accounting Office and the elected President of both the American Evaluation Association and the Knowledge Utilization Society, she brought a great deal of relevant context knowledge and expertise to this work.

She found that the strategy suggested by the National Academy's Committee was just the tip of a pyramid of possible consensus strategies that OERI should consider. This pyramid showing the range of "Types of Consensus Strategies" is shown in Figure 1. The pyramid has a broad base with many systematic decision making strategies each of which includes at least some consensus related activity (such as group discussions or syntheses of comments from various reviewers) and a top point with a few model consensus panel approaches which require extensive participant agreement and rigorous examinations of evidence. To show the dual focus of this report on R&D agenda development and R&D dissemination decisions,

Figure 1.—Types of Consensus Strategies



this pyramid is divided in half vertically. In this report, Dr. Datta has emphasized the top and middle levels of the pyramid. The top level consensus panel models are characterized by the relatively high cost and time consuming work of the Office of Medical Applications and Research in the U.S. Department of Health and Human Services. Next come the more varied consensus-building strategies that involve procedures to secure some agreement on what's best to do or recommend. The numerous group decision-making strategies at the base of the pyramid include techniques such as preparing commission reports, using judicial type processes, sharing drafts of syntheses for review and revision, and soliciting and aggregating input from stakeholders. Users of the numerous democratic approaches to government decision-making at the bottom of the pyramid should also benefit from insights gained from the consensus intensive strategies in the top two sections of the pyramid.

While Dr. Datta found merit in the Committee's overall recommendation to use consensus development processes, she recommends that OERI try the consensus panel strategy specified in the National Academy of Sciences report initially only on a trial basis. This is because the necessary features of effective consensus panels in health research and practice probably can not be transferred or adapted to educational research and practice. Despite this cautious recommendation, Dr. Datta finds a great deal of potential benefit in OERI's appropriate use of consensus building efforts. *A Matter of Consensus* describes the strengths and weaknesses of the relatively expensive consensus panels as well as many less expensive consensus development approaches as they are used for both R&D agenda planning and federal level decisions on what R&D is worth disseminating.

In the main section of the report Dr. Datta examines ten questions and answers about consensus strategies and provides general and specific recommendations on issues related to the important federal roles of planning R&D agendas and providing responsible information to consumers on their best options. This section describes many of the complexities and possibilities inherent in national consensus development work.

However, readers who want to benefit from the full depth of Dr. Datta's insights, should not overlook the additional analytic information provided in Appendices IV-VII. Appendices IV and V contain 10 descriptive analyses of consensus panels and consensus development processes. Appendices VI and VII examine some specific experiences of OERI and other agencies in developing R&D based consensus on topics such as how to teach reading.

Dr. Datta's broad scope and indepth analysis of consensus strategies will be particularly valuable in guiding the long term development of R&D agendas in the five new OERI Institutes and in implementing OERI's immediate and continuous mandate to identify and share the best education products, programs, and practices. However, since many of the consensus strategy examples come from outside of education, and since the principles are appropriate for many federal challenges, *A Matter of Consensus* should be a useful tool for all types of federal R&D-based social science decision-making. This report can become an

even more valuable decision analysis tool if it is used with two other OERI led efforts to improve R&D planning and dissemination. They are: "A National Agenda for Education Research and Development: Why not an A-Plus?" by OERI staff, Elizabeth Demarest and Sheryl Stein (Feb. 1994 draft) and a special feature of *Evaluation and Program Planning* on "Sharing the Best: Finding Better Ways for the Federal Government to Use Evaluation to Guide the Dissemination of Promising and Exemplary Education Solutions" by Department of Education staff and others (Klein, 1993).

You are invited to view this report as a working document where your additional information on consensus strategies and feedback on the current information will be welcome as an example of how OERI is trying to "practice what it preaches" by obtaining expert feedback from concerned consumers to arrive at increased well-informed consensus on what strategies it should support.

A MATTER OF CONSENSUS

By Dr. Lois-ellin Datta
Datta Analysis, Waikoloa, Hawaii

TEN QUESTIONS AND ANSWERS ABOUT CONSENSUS PANELS

1. WHAT ARE CONSENSUS PANELS?

Consensus means a coming to agreement. According to the National Institutes of Health, a consensus development conference--the archetype of this approach---

> "..evaluates the available scientific information on a biomedical technology and develops a consensus statement that advances understanding of the technology or issue in question (assessment) and that will be useful to health professionals and the public at large (transfer). (NIH, undated, p. 1)

If the key elements are **evaluating available information** and **developing a consensus statement**, then consensus-like processes range from procedures for reaching unanimity (such as juries, the Society of Friends sense of the meeting process, the College of Cardinals selection of a new Pope, and the panels developing national standards for educational evaluation, mathematics and science) to procedures for obtaining broad input to find out where the field as a whole is moving or wants to go (such as conferences on motivation for learning and sending draft research agendas out for review by diverse stakeholders.)

As Table 1 indicates, these take many forms; appear under many names; and have varying degrees of common elements. It seems useful to talk about two types:

> **Consensus Building Approaches** aim at reaching maximum common ground on topics through a process of bringing the views of diverse people to bear on relevant information and developing a sense-of-the-field statement.
>
> **Consensus Panels** aim at reaching maximum common ground on topics that are significant and probably controversial through a process of bringing diverse people together to examine the evidence and craft as unanimous a statement as possible making specific recommendations on the topic.

The Environmental Protection Agency (EPA), in the context of negotiated rule-making, has a useful chart (page 3 of the EPA Appendix in this report) which illustrates an elegantly conceptualized spectrum of consultation and consensus building activities, ranging from one-time public hearings for information exchange to consensus oriented negotiation in the regulatory process. The figure also shows at what point, in the EPA application, federal regulations regarding advisory committees (Federal Advisory Committee Act) and public discussions (Administrative Procedures Act) must be followed (EPA, March 4, 1992).

TABLE 1: SOME EXAMPLES OF THE RANGE OF FEATURES IN CONSENSUS BUILDING APPROACHES AND CONSENSUS PANELS

DIMENSION	RANGE		
Charge	Broad	to	Highly specific
Stakes	Awareness	to	Force of law
Topic selection	Sponsor Interest	to	Tight criteria
Basis for selection of panelists	Random	to	Narrowly Selective
Literature reviews	None	to	Extensive
Source of evidence	Experience	to	Gold-standard research
Agreement required	None	to	Unanimity
Sponsor clearances	None	to	Extensive
Time spent in panel interactions prior to statement (in person plus other means)	2-5 days	to	Two years

2. WHO USES CONSENSUS BUILDING APPROACHES AND CONSENSUS PANELS?

Just about all research organizations one can think of, including OERI/ED, seem to use **consensus building approaches**, usually liberally.

Consensus panels, as defined, are scarcer but the examples include:

- local governments (for example, Princeton, NJ)
- non-profit foundations (for example, the Jefferson Center for New Democratic Processes)
- Congressional agencies (for example, the Office of Technology Assessment)
- Federal agencies (for example, the Environmental Protection Agency; the Department of Health and Human Services; the Education Department)

Some of these instances are one or two time events (the Princeton, N.J. panel; the NIE/ED convening process). Other instances are major institutional strategies with well-established procedures, many examples of use, and evaluations of their effectiveness, such as those of the National Institutes of Health consensus development program.

3. WHAT, SPECIFICALLY, HAVE THE CONSENSUS BUILDING APPROACHES BEEN USED FOR?

Consensus building approaches, commissions, and task forces are familiar to OERI/ED. To mention only a few instances, they have been have been used to:

- Develop research agendas in reading, teaching, and cognitive skills for NIE/ED; priorities for laboratory and center competitions for OERI/ED; and research agendas in educational statistics for OERI/ED

- Prepare statements on implications of research for practice in reading and student motivation for OERI/ED

- Analyze broad policy issues and craft policy statements such as the National Governors' Association and others in creating the Year 2000 national educational goals

- Establish a national consensus on what's wrong with education and what should be done about it, such as the commission whose work culminated in "A Nation At Risk"

- Review the state-of-knowledge on topics such as student achievement testing which may include recommendations for policy, practice and further research (National Academy of Science)

- Examine the state of an area such as educational research and educational statistics (National Research Council)

- Preparing national statements on what research says about issues, goals and objectives (Review of Research on Achieving the Nation's Readiness Goal, OERI, 1993)

Many other agencies, by choice or as required by statute to assure broad participation and expert advice, use consensus building approaches, commissions, and advisory boards. The statements of purpose for many of these have, as their official reasons for spending time and money on the efforts, assuring diverse input and efficiently tapping into the nation's greatest expertise. However, squawk reduction through participation, negotiation and consensus development can be a happy by-product.

Currently there are about 1,200 such standing panels in about 50 research agencies.

As one example, the Environmental Protection Agency has 22 separate advisory groups whose "collective scientific knowledge, independent perspectives, and outreach to key constituencies enable (EPA) to make better decisions, which will protect human health and improvement the quality of human life." (EPA, 1992, p. i). These include the Council on Clear Air Act Compliance Analysis, the Lawn Care Pesticide Advisory Committee, and the Environmental Statistics Advisory Committee. These Committees not only review what the agency does and how it does it, they are often charged by Congress with preparing independent reports that include analysis of the state of knowledge and practice.

Of particular interest is the EPA Negotiated Rulemaking Process, developed as an additional approach for the agency in carrying out its rulemaking responsibilities. The EPA experience is a reminder that proximity can exacerbate disagreements as well as resolve them. About half of EPA early (1983-1987) negotiated consensus panels did not achieve closure and in one case, a key group walked out. EPA understandably includes among its criteria for topic selection analysis of the risk that the consensus effort will backfire (EPA, 1993).

4. WHAT, SPECIFICALLY, HAVE THE CONSENSUS PANELS BEEN USED FOR?

The Education Department again is no stranger to consensus panels, both in the one-two time use form and in the standing form.

- The NIE/ED convening process brought academic and craft knowledge to bear in two instances: the District of Columbia Public Schools analysis of its competency-based curriculum and the Mississippi state plan for teacher certification and school accreditation (Klein, Gold and Stalford, 1986).

- The National Institute on Disability and Rehabilitation Research began a Consensus Validation Conference Program in 1992. The purpose is "to evaluate and synthesize available scientific information and improve dissemination of findings from rehabilitation research." (NIDRR, undated.) About eight conferences will be held by 1995 on topics such as strategies to secure and maintain employment for people with long-term mental illness and augmentative and alternative communication interventions. The NIDRR experience is surely a rich lode from which OERI could draw.

- Many of the Department's recognition programs---of teachers, of schools, of programs---involve (a) a fairly broad-based group of reviewers, (b) some evidence of quality or effectiveness, and (c) an effort to reach agreement on the decisions. The Program Effectiveness Panel is a sort of hybrid: consensus-panel-like in an emphasis on evidence of the effectiveness of innovations and in seeking diversity among panel members, but not consensus-panel-like in that decisions are made by majority vote and, at present, panels do not meet in person (Ralph and Dwyer, 1988; Stalford, personal communication, 1994).

- A clearer example may be Department support of the various standard-setting efforts in education, such as the educational evaluation standards, the standards for mathematics and science, and the standards for literacy. These efforts represent a consensus among diverse stakeholders, involve face-to-face discussions, bring up controversial matters and make specific recommendations. They are not consensus-panel-like in that clarification of issues through scientific and technical issues may play a smaller role than "clinical experience" and values.

There are, however, organizations with more extensive experience with consensus panels and more diverse applications. As examples,

- In the non-profit foundation sector, since 1974, the Jefferson Center for New Democratic Process has run Citizens' Juries on topics such as health care reform and the federal budget (Crosby, 1993). After a topic has been selected

(by the Center or at the request of federal, state or local governments), staff develop a framework of basic decisions and bring together relevant evidence and witnesses, a random sample of citizens is drawn, 2 to 4 days of public hearings and discussion are held, and the Citizens' Jury makes its decision. Skilled involvement of mass media usually brings much attention to the issues as framed and the jury's statements, as representatives of the people. **(Policy decisions example)**

- In the non-profit organization sector, the American Psychological Society has begun a broad-based, ambitious **research agenda development** consensus panel process. The long-term aim is to get many research associations to coalesce around a common agenda, expressed in a vision that can ignite Congressional enthusiasm. The first product of this multi-year effort, "The Human Capital Initiative," was developed through two summit conferences involving over 70 organizations including the American Educational Research Association. The effort took about two years (APS,1992)

- Also in the non-profit organization sector, the National Stroke Association turned to a consensus panel process to develop a statement indicating that research shows the aggressive emergency evaluation and treatment of stroke notably improves recovery. The panel was composed of ten top national stroke authorities; the statement has been endorsed by the majority of neurology-related medical associations in the world; over 100,000 copies have been distributed to emergency room physicians and nurses---in addition to press conferences and training kit development. (National Stroke Association, 1993; Houser, personal communication, 1993). **(Research into practice example).**

- Among congressional agencies, the Office of Technology Assessment standard operating procedure for preparing their reports to Congress is the extended consensus panel **(state of knowledge example)**.

- Turning to executive branch agencies, the Department of Health and Human Services Administration for Children and Families established a consensus panel for **research agenda development** when enough funds became available to restart their research program.

- The Environmental Protection Agency has a well-regarded Negotiated Rulemaking Process **(regulatory example)**. EPA has found that "In the right situations, negotiated rule-making can produce proposed rules that meet statutory requirements but are more pragmatic than proposals EPA would be likely to develop on its own and may produce better environmental results;...and are more likely (than conventionally developed rules) to be accepted by the affected industries and other interested parties involved in developing them." (EPA, 1987; Kelly, personal communication, 1993). In addition, EPA uses a consensus process for **policy development**. The agency has a detailed chart for helping decide what type of consensus strategy would be most appropriate for what type of regulatory

situation. (See attachment). This is a matter of high priority for an organization with about 250 regulations in process at any time.

- Perhaps the best-known, possibly the best-documented, and probably the most thoroughly-evaluated use of consensus panels is the National Institutes of Health Office of Medical Applications of Research's (OMAR's) consensus development conferences. Since 1977, over 90 conferences have been held, on topics such as CEA as a cancer marker, dental sealants in the prevention of tooth decay, travellers' diarrhea, platelet transfusion therapy, cochlear implants, urinary incontinence, acoustic neuroma, impotence, and early identification of hearing impairment in young adults. The aim is to go from **research into practice** where there is sound evidence of the efficacy and safety of new medical technologies and approaches (NIH/OMAR, undated).

  The process has evolved to fit closely into the NIH structure and the federal/health care system relationship. OMAR is in the Office of the NIH Director and, as a component of this Office, is charged with dissemination. It has money (to pay for the conferences and dissemination); it has staff (for technical assistance oversight to assure the integrity of the process); and it has influence (in outreach to the media and key journals and because a panel statement on the efficacy and safety of a procedure can enable health care providers in the massive Medicare, Medicaid and Champus systems to be reimbursed for using the approach).

  NIH emphasizes results so the Directors of each Institute have many incentives for proposing conference topics to OMAR. The effort is a close partnership. The sponsor Institute is responsible for assigning senior staff with excellent scientific knowledge to guide all technical aspects of the effort. OMAR provides know-how in the management of the conference process, logistic support through a standing contract, and assistance in disseminating the consensus panel statement. The statement is the panel's own; given directly and immediately in a press conference on the last day of the conference; and published without agency clearances very shortly thereafter.

  This is, obviously, neither a quick nor a cheap process--appropriately, given the stakes! About two years are required, and each panel costs, on the average, between $150,000 and $160,000 according to agency officials whose estimates may not include all the cost items such as the value of staff time included in some of my other figures.

  Third-party evaluations of the panel process in its earlier form while recognizing the strengths, saw the panels as a work-in-progress and identified improvements OMAR officials state are now in place to increase outreach and impact (Vinokur et al., 1985.Wortman et al., 1988; Kanouse et al., 1989). These evaluations, which were actually carried out in the early 1980s, were technically sound and indeed,

state-of-the-art, for their time. Third-party evaluations of the more recent CDP as it has evolved are not available although the 1994 HHS/Inspector General's "replication" of one component of the Kanouse et al. study was an independent analysis showing progress with regard to outreach through the Continuing Medical Education programs but also room for improvement. OMAR also conducts formative evaluations in-house.

- A new Public Health Service (PHS) use of consensus panels for **applying research and clinical experience to practice** is being developed under the NIH's Agency for Health Care Policy and Research (AHCPR). AHCPR was authorized in 1989 and as such, has issued about seven consensus panel statements. However, for over a decade, work has been underway on consensus statements on appropriate and necessary medical treatment. Several variants are being tested. One variant (RANDs), uses a literature review to identify conditions where a given treatment has been found useful. The panel is asked to validate whether these conditions so match clinical experience as to be a standard for what is necessary and appropriate use of the treatment.

  The AHCPR research on this method may be particularly informative for education, since a driving need for clinical expert panels that the scientific literature is not as complete or as high a quality as usually sought by OMAR (Shekelle, 1992).

Neither the OMAR nor the AHCPR consensus panel statements have always been received with acclaim, despite the careful process. As statements rely more on clinical experience and less on gold-standard research, they may be more vulnerable to criticism and rejection by some health providers. For example, the 1993 AHCPR consensus statement on treatment of depression was severely critique by the American Psychological Association (among other things, for neglect of psychologists and psychological approaches and alleged bias toward primary care physicians and pharmacological approaches). AHCPR is "revisiting" the depression guidelines in 1994, before the usual 18-36 months allowed for new findings to come in (Clinton et al., 1994; Schulberg and Rush, 1994; Munoz et al., 1994). (AHCPR provides three different documents for each panel: the scientific bases and literature reviews; a health care provider guide; and a patient guide (in English and Spanish versions. The tumult centers on the provider and patient guides.)

On the other hand, the 1994 AHCPR Guidelines urging physicians to treat cancer pain urgently, early, and diligently without worrying about patient addictions seem to be well-received and widely applauded. These new guidelines emphasize physician responsibilities and a sequence of options starting with mild painkillers like aspirin and including unlimited dosages of morphine and other powerful drugs as needed.

5. HOW MUCH DO THESE PANELS COST?

Hard information on panel costs is hard to come by. Little or nothing is published. Interviewees who are splendid sources of purposes, products, and results either don't have detailed cost analyses or feel (in the non-federal sector) this is privileged information.

Further, the basis of what cost data could be obtained often was apples, oranges and orangutans among panels. For example, an effort described as costing $20,000 "on a shoe-string" meant funds available for small travel grants. When the value of donated time, travel costs subsidized by others, staff time, indirect use of other grant funds, and such were included, my estimate of costs came close to that for comparable fully-funded panels.

The art of managing consensus-building panels in general is passed on through apprenticeship, craft knowledge and retrospective analyses. The costs and time depend on the number of persons involved; the magnitude of the topic; and the modesty (or majesty) with which the reports and statements are disseminated. About $100,000 and 12 months seems to be modal for a one-time conference involving perhaps 20 persons. NIH/OMAR indicates that its consensus development conferences cost on the average between $150,000 and $160,000.

A standing panel meeting many times with various missions could cost more. For example, the EPA Acid Rain Advisory Committee involves 44 members, (public utility commissioners, state air pollution control officers, utility executives, environmentalists, consumer advocates, pollution control representatives, academicians, and individuals representing coal and gas interests). In FY 1991, the full committee met six times in a seven month period and held five subcommittee meetings to "allow for an interplay of diverse opinions and positions by various individuals and groups helping to generate workable solutions to potentially problematic issues in the regulations." (EPA, 1992, p.7) If a logistics contract to manage travel, meeting, distribution of advance materials and such was awarded to a for-profit firm, and the direct and indirect costs of this activity, including contract management, liaison by agency staff, and report production are considered, I estimate that an effort of this magnitude could run above $300,000.

In general, there's a highly skewed distribution. My estimate for most panels (not including costs of post-panel dissemination) is about $100,000 per consensus panel, pretty much across different applications. The high end can get quite high, however, when many meetings are required, when large numbers of stakeholders must be involved, and particularly when dissemination gets into video for the media, training kits, and hundreds of thousands of statements distributed.

RAND is experimenting with cheaper and possibly faster methods, such as Delphi techniques, in controlled tests of results of different panels drawn from the same pool examining the same topic. The studies compare the extended face-to-face, mediated discussion method with other consensus panel techniques. Computer-assisted conferencing and other telecommunications applications may have many of the benefits of face-to-face meetings and would seem to have good potential for cost-reduction, although it seems likely these can not substitute for some face-to-face discussions.

6. HOW FAST?

Based on conversations and experience, rather than systematic comparisons, my estimate is that consensus panels are not very fast, compared to say, giving a bright, task-oriented, experienced and unbiased person a free rein to read the literature and, without having to seek consensus, reviews, or clearances, prepare a statement.

Further, since literature reviews generally are taken as the beginning of the consensus conference process, the consensus building approaches and consensus panels almost by definition take considerably longer than such a literature review alone. However, such comparisons are slippery since the goals of a consensus process go beyond simply getting information together to creating a momentum, an irresistible surge or consolidation of opinion.

Again, hard data are hard to come by. What is included in the timeline can vary from estimate to estimate. Taking a common start point as initiation of the process for selecting the topic and a common end point as a statement-ready-for-announcement, my estimate is a skewed distribution with about 12 to 18 months required for an **experienced** organization to run a formal consensus panel. (Some approaches are faster; the Citizens' Juries can be put together more quickly; the NIE Convening Process for the District of Columbia Public Schools took about 6 months from initiation to completion. Some, such as OMAR, run slower.)

This about 12 to 18 month estimate is for a panel making an independent statement **that does not require sponsoring agency clearance, review, revision, "proof-reading", "editing", or vetting.** The time can get much longer if the statement is really an institutional position, and is edited, reviewed, negotiated, vetoed within the agency after the panel finishes its deliberation. Presumably, there is a fairly fast learning curve to help an organization be this speedy. However, each panel seems so adapted to its purposes, organizational structure and relations between the organization and the stakeholders to be reached that the processes clearly aren't being Xeroxed. Time would be needed to design a locally appropriate process, try it out, revise, and try again. **The consensus panel process seems to be a living experience, evolving and being fine-tuned even within experienced agencies such as OMAR, and certainly needing development, try-out, and revision again for newly-adopting agencies.**

There are other considerations affecting time. The NIH/OMAR process may take from 1 to 2 or more years (not to actually hold the panels but from the start to end point as defined above). Assurance of unimpeachable scientific integrity, thoroughness of review of the evidence, and other quality concerns have time as well as direct cost trade-offs. One of the reasons the National Stroke Association held its consensus conference on the emergency-room assessment of stroke damage and initiation of therapy is that NIH officials advised them that it could be several years before knowledge that could save lives went through all aspects of the OMAR process. (Houser, personal communication, 1993.)

7. WHY GO TO ALL THE TROUBLE? HOW EFFECTIVE ARE THE CONSENSUS BUILDING APPROACHES AND THE CONSENSUS PANELS FOR RESEARCH AGENDA SETTING?

Consensus building approaches or consensus panels for research agenda setting are often used when a research program is to be initiated: the NIE reading research example, the Administration for Children and Families research program restarting example; the American Psychological Society national research agenda initiative. Either money is in the offing or the consensus plus a plan with an exciting vision may attract enough money to be worth the effort. When the panel may cost a goodly percent funds available for research, however, it seems to take a strong stomach to invest in planning.

There appear to have been no readily available systematic evaluations of the cost/benefits of consensus building approaches or consensus panels in research agenda setting. Such analyses might look, for example, at research agendas on the same topics developed through different methods. Descriptive analyses might include the nature, content and number of the substantive recommendations; their linkages to prior research, to practice, to emerging issues; and the clarity with which they're presented. Results-oriented analyses might look at initial stakeholder response to the agenda, funding source response, extent to which proposals, awards, papers, and books eventually reflected the agenda. and the productivity (or lack of it) of research initiated. Efficiency analyses might look at costs and time for agenda preparation, and extent of modification from panels to announcements. Lacking even a descriptive comparison, the fall-back is more qualitative judgments.

The NIE conferences generally have gotten great press and have led to well-received state-of-knowledge books as well as guiding research not only in the Education Department but also in other agencies such as the Office of Naval Research and the National Institutes of Health. Kraut (personal communication, 1993) has a sheaf of clippings showing the "The Human Capital Initiative" has garnered Congressional enthusiasm and considerable agency attention. In the future, the American Psychological Society could have valuable interrupted time series data on the leverage of its conferences. NSF, for example, is dedicating about 20% of all its increase in behavioral science funds to the relevant APS agenda topics.

The reverse side of the coin is perhaps clearer. Case studies and histories of what has gone wrong in federal research agencies show failure to use consensus processes to set research agendas can be dangerous for their health. The issue is what degree of consensus building, through how elaborate a process, does the agency feel will maximize support and minimize funds diverted from direct funding. One approach is to build on the second application of consensus panels. Where literature is synthesized to inform practice, identification of knowledge gaps can inform agenda development, a point which has been made previously by many writers. That is, the end point of going from research into practice and its by-product, knowing where research is and is not robust, routinely could be among the starting points for a consensus building process on a disciplined research agenda.

8. WHY GO TO ALL THE TROUBLE? HOW EFFECTIVE ARE THE CONSENSUS BUILDING APPROACHES AND CONSENSUS PANELS FOR DISSEMINATION?

The costs/benefits of the dissemination applications of consensus building and consensus panels seem to depend on quality of the panel statement (Kanouse et al., 1989); on the extent of dissemination (Kanouse et al., 1989; Jefferson Center personal communication, 1993; Blaisdell, personal communication, 1993); and probably on the consequences for paying attention to the statement.

Agreement seems high that consensus panel statements have greater credibility than a paper by a single individual, however well reviewed. There also seems to be agreement that the panel process, when carefully and scrupulously conducted, gives a better result in thorough examination of all relevant evidence through many lenses. But absent something like the Medicare system or malpractice suits, the statement may join the many pronouncements beamed to busy practitioners.

More specifically, the Kanouse et al. (1989) evaluation of the early 1978 to 1983 OMAR consensus statements found:

- variability in statement clarity ranging from near- impossibility in determining what one would do to be in compliance to certainty about being in compliance would entail (OMAR has since aimed for practice-oriented, didactic, do-this statements. Examples of consensus statements from 1979 to 1993 are included in the attachments to this paper, and clearly show improvements in clarity, readability, and information given.)

- about a three-month window of maximum attention in the patient-oriented and popular media to the statements (OMAR has since aimed for good media attention on the last day of the panels)

- much greater recognition of the message than the messenger; physicians were more likely to know about the substance of the recommendations than to connect them to the NIH Consensus Development Program.

- a slower change in citations in the research journals but the statements are sort of catching an already-rising curve of attention; also, citations didn't always mean agreement with the statement.

- much greater impacts on physician awareness of the statement and change in practice when the messages were delivered through Continuing Medical Education (CME) programs and key journals such as the Journal of the American Medical Association. (OMAR has since expanded CME outreach , and sends statements out immediately for publication in key journals. Actual publication may take, however, 5 to 6 months in journals such as JAMA.)

- much greater impact on actual practice when the topic wasn't flogging a dead pony (advocating a change already wide-spread) and when it was dealing with a highly significant topic with "user readiness".

This 1989 study of the early 1978-1983 panels found that between 15% and 30% of all physicians were aware of the statement, with higher percentages for certain topics (such as C-sections) and for certain medical specialties. It seems likely that in 1994, the OMAR statements are now more an established part of the physician landscape, and that their outreach and impact are quite significant. In addition to disseminating research information to improve practice, the OMAR consensus panel reports address implications for further research.

The glass, at present, may be seen by Department of Health and Human Services senior officials as half-full in terms of progress since the earlier period but somewhat empty in terms of room for improvement. For example, 52 percent of a population survey of medical school department chairs "reported having sponsored a continuing education activity that addressed the findings of a recent NIH consensus panel conference." (DHS, 1994, p. I) The Department regarded this as too low a percentage and the report suggests ways to strengthen outreach and deal with some specific concerns (particularly about panel impartiality, qualifications and time for reflection). However, this study probably underestimated the outreach of the recommendations because the questionnaire focussed on identification of the NIH "brand name", which the 1989 study found was less salient than whether the specific recommendations were known.

9. IS THERE ANYTHING HERE FOR OERI/ED?

Yes and no; it depends on the application.

In the yes column, the consensus building and consensus panel processes seem useful for the research agenda development. How much, and with which degree of unanimity, would depend on resources available, whether the boat has already left the dock, and how much OERI wants to invest in the future. Looking in some depth at joining the APS effort (if this is not already in process) might be a relatively high benefit, moderate cost way for OERI to build for the future in research agenda setting. Put another way, OERI would miss leveraging an important opportunity if it doesn't commission an APS-like effort linked to the Human Capital Initiative category of Schooling and Literacy.

With regard to other applications, OERI itself doesn't issue many regulations, but might serve as a partner to other offices within the Department in trying out the negotiated regulatory process. The Environmental Protection Agency has some excellent materials that could be starting points for selecting which regulations might be good candidates for the negotiated approach and which would not. For certain key procurements, such as the Laboratories and Centers awards, OERI might look to the EPA model for itself.

Further, OERI also might be a partner to other offices and top leadership in trying out the consensus panel approach in areas such as policy development.

Also in the yes column is the **consensus building approach** for OERI's many research-into-practice efforts. For example, capturing in a guidebook craft knowledge about pitfalls in managing research syntheses and how to avoid them could be well-worth while for an organization that puts out as many publications as OERI does. There are certainly a lot of pitfalls, ranging from selection of prima donnas (and dons) to dealing with high emotions. And there seems to be a lot of craft knowledge out there, that could be combined with formal research on negotiations and group dynamics (Anderson, personal communication, 1993; Backer, personal communication, 1993; Backer, 1991, Wortman et al., 1982) for OERI project officers, writers, and facilitators. Such a book could be a companion to the many analyses of technical aspects of achieving good synthesis quality (Hedges, 1984; Cooper and Harris, 1993; Ward and Reed, 1983).

10. AND IN THE NO COLUMN

In the no column: adoption full-bore of the NIH-type of **consensus panel** for dissemination statements or other ways of going from solutions/research into practice.

This conclusion is different from that of others, such as the NRC panel on educational research who advocate use of the NIH-type of consensus panels for OERI. (See also Havelock, 1986 for analysis of issues.)

Why so cautious? Because the recommendations probably were made as one way to achieve Respect, Attention and Impact for the results of educational research comparable to that of medical research, both for the good of the order in future research funding and to accelerate the improvement of education in this country.

The issue is a classic one in dissemination: there is an innovation; it looks effective; will it transfer? That is, is the education research context sufficiently similar to that of medical research to expect good transferability?

My analysis of the similarities and differences in contexts certainly shows some similarities, but too many dissimilarities to recommend full speed ahead.

A summary of some of the OMAR criteria for selecting topics may illustrate the reasons why NIH procedures and NIH results may not be easily obtainable for OERI. A topic must meet the following NIH selection criteria:

- it should have broad, significant public health importance
- controversy or unresolved issues should surround the biomedical aspects of the topic that could be clarified by the consensus approach
- it must have an adequately defined and available base of scientific information from which to answer the conference questions and to resolve the controversies
- it must be amenable to clarification on technical grounds and the outcome should not be dependent mainly on the subjective judgments or values of panelists

Comparing educational research to these criteria, first the educational knowledge pool may be a mile wide but an inch deep in many places. NIH currently spends over $5 billion annually on health research of which about $150 million goes to clinical trials alone. In comparison, OERI barely has $2 million annually for unsolicited research proposals, and as a rough estimate, maybe $20 million net of overhead and other activities annually for basic educational research through the centers as a direct research expenditure. In an effort to cover all bases, OERI has chosen to or been directed by Congress to diversify across almost 20 educational topics. Further, and perhaps most crucially, there is nothing comparable to a clinical trials program for educational research and development.

Also unclear is the size of the pool of educational issues that are significant, controversial, and amenable to clarification on technical grounds rather than subjective judgments or ideology. For example, could OERI make a statement about student ability grouping for instructional purposes (homogeneous versus heterogenous student assignments) based primarily on top-quality research on outcomes such as student academic learning, personal-social development or teacher efficiency? Is this topic constrained by legislation (integration of students with handicapping conditions) and beliefs about much larger social goals as well as a possibly limited pool of research? Meta-analytic techniques for knowledge synthesis may help extract the greatest possible information from the educational knowledge pools, but it seems likely that consensus procedures bringing both research and craft knowledge/clinical experience to bear, such as AHCPR's clinical practice guidelines, may best mesh with the state of educational research.

In addition, the NIH process excludes (on the basis of conflict of interest) from the actual panel the persons who have themselves done the relevant research; and utmost care is taken to maintain the diversity, integrity and independence of the panel, including no NIH clearance, editing or other prior review for panel statements. Would this be politically possible for OERI? It apparently has not been possible for NIDRR.

Finally, the NIH consensus panel program is embedded in an infra-structure emphasizing continuing medical education and explicit standards of practice as well as attention to research. **There are liabilities for not meeting standards in malpractice suites. And in payment structures, there are incentives for adhering to statements of necessary and appropriate practice.** AHCPR was launched as part of the health cost containment efforts. The health system structure hardly works perfectly but is more performance-oriented than education and there are federal systems (in Medicare, CHAMPUS) for putting research based guidelines into practice. (See also Table 2).

HOWEVER, THERE ARE REASONS FOR GOING AHEAD on a pilot or experimental basis, with the expectation that considerable re-design, re-structuring and acculturation may be needed. First, there may be areas rich in research that haven't already had umpteen reviews of the literature and OERI/ED research-into-practice guides, thus making a consensus panel feasible and worthwhile. Second, the world of medical research is not all perfect and neat while educational research is all imperfect and messy. OMAR persevered in finding a way through the obstacles, AHCPR is learning, and OERI should at least give it a good try, since the potential rewards are indeed high. Third, the "flagship" work of the NIDRR panels within the Department, the potential similarity of the relationship between OMAR and the NIH Institutes, and the OERI Office of Dissemination and the OERI Institutes plus the clean slate of radical changes argue for a pilot-basis trial. If this is a direction OERI chooses, both AHCPR and OMAR leadership would be only a few subway stops away for direct, detailed consultation.

TABLE 2: "GO" DECISION INDICATORS FOR USING CONSENSUS PANELS FOR DISSEMINATION

- There are relatively few decisions to be made
- There is considerable time (12 to 24 month from initiation to public statement) available
- Money and/or internal staff are available (about $100,000 per panel seems to be a rough median of direct costs)
- There is a major issue such as whether research shows a common practice is bad and should be discontinued or whether a new product, product or policy is highly effective and should replace current practice
- The issue is solution- focused such as whether assisted methods improve communication enough to be worthwhile, relative to risks and costs of alternative procedures, in general practice
- Good information is available about current practice so time isn't wasted urging a change that already has been made
- The political stakes are low (or can be kept far away) so the technical/scientific stakes can predominate. Decisions can be made on evidence, not political values
- There is enough empirical evidence of a decent quality to be worth synthesizing but the outcomes are not already so widely accepted that the results would be ho-hum and the topic hasn't already been research-synthesized to saturation
- The demands of the panel process are taken seriously. The extensive pre-preparation, the top quality knowledge syntheses, the attention to the dynamic of the panels and what is required for sound panel results are taken into account when the consensus panel is designed.
- The integrity of the panel process can be protected. Dissemination in agencies such as NIH is a partly centralized function. The panel process integrity is protected by the central office but the actual panels are substantively guided through the separate institutes, trying for the best of both worlds, in a squeaky-clean, you-can- trust-the-integrity-of-the-process setting.

RECOMMENDATIONS

1. CLEARING THE DEFINITIONAL UNDERBRUSH:

The U.S. Department of Education/OERI should consider using a variety of consultative and consensus building activities, but bring some definitional (and quality) order in to the use of terms. Each would have a clearly-stated methodology for all aspects of the panel process, from topic selection through reviews (or not) prior to public release (see OMAR example.) Each would be monitored or audited periodically to assure the integrity of the process and for formative purposes.

These activities might include (but not be limited to):

- consensus panels (reserved for procedures with high stakes and high scientific integrity on nationally significant questions, released without agency clearance; unanimity sought among relevant diverse perspectives on solution-oriented programs, practices, products)

- consensus building approaches (maximum agreement sought among relevant diverse perspectives)

- validation panels (panel discussions of positions drafted by staff and/or non-federal experts, with agency clearances; agreement sought among relevant diverse perspectives)

- convening panels (bringing craft/research knowledge to the resolution of state/local educational issues or disputes; agreement sought among relevant diverse perspectives)

- effectiveness review panels (panel reviews of the quality/effectiveness of specific programs, practices and products relying on scientific/technical information and involving some discussions; agreement sought among relevant diverse perspectives)

2. TRY IT! SOME PROMISING APPLICATIONS:

There are uncertainties about applicability for OERI of using the consensus panel processes for some high stake decisions in dissemination on a pilot basis, however, the approach is worth trying. Further, there are other options involving panels that are not now prevalent in OERI but are used by other organizations that are worth OERI consideration as part of its utilization mission. These include:

- Long-range policy formulation, before political imperatives limit considerations: getting ahead of the curve using research findings

- Conflict mediation in highly selected instances

- Through TACs, for decisions which are moderate stakes and with some generalizability so cost can be lower on a per/use basis for findings

3. THE COMPLIMENTARY OF DISSEMINATION AND RESEARCH EXPERTISE: Assume ED/OERI has qualified staff, enough money, and a willingness to see how the consensus panel approach could be adapted to education. In that case, to help assure the quality and integrity of the process, consensus panels for dissemination should be centrally managed across ED programs in something like the NIH OMAR/Institute relationship.

That is, each Institute Director could apply each year for ED/OD (Office of Dissemination) selection for consensus panels on topics for the specific Institutes. ED/OD would review these applications and using criteria such as those suggested, select the most promising topics. ED/OD funds would be made available for the selected panels of each Institute, with assignment of ED/OD staff to provide technical assistance and to monitor the integrity and quality of the panel process. ED/OD would be responsible for developing the detailed manuals and guidelines, based on state-of-the-art experience. The ED/OD also would have funds to commission independent research and evaluation on crucial topics needed to improve the ED/OD Consensus Panel approach.

To promote maximum scientific integrity, the NIH/OMAR safeguards against political interference and for statement authority should be followed. They've learned a lot about this aspect, although other aspects of the NIH/OMAR approach, such what research designs would substitute for the randomized clinical trials gold standard, would need modification for ED. ACHPR's process could provide a useful starting point.

APPENDIX I: METHODOLOGY

Methodologically, the aim of this paper was to describe the range of consensus panel and consensus building panels for research agenda development and dissemination of research findings to improve educational practice. In the time available, comprehensiveness (in the sense of identifying every single use of such panels) was both impractical and of less priority than assurance that highly significant applications were captured. Further, the aim was to get as much information as possible about panel processes, costs, strengths, weaknesses, uncertainties, particularly high-quality third-party evaluations of the panels' effectiveness. Lastly, the aim was to understand issues of transfer, reproductibility, adaptation that might limit the value in the education community of an approach that worked niftily in another setting.

Three approaches were used in collecting information for this paper: (1) computer literature searches by OERI staff, (2) follow-up interviews in various agencies on the uses of consensus panels cited in the National Research Council report, and (3) networking with experts in the field, including publication of a "seriously seeking consensus panels" announcement in the National Dissemination Association newsletter and obtaining the names of key persons known to experts in OERI such as Susan Klein, Charles Stalford, and those who attended the November 1993 meeting. A fourth approach was serendipity through both media (including TV, magazines and newspapers) and personal contacts.

The initial goal was to obtain written information, describing the panels and providing details on processes, costs, and results. Particularly sought were any third-party studies or evaluations. There weren't many such studies except in the health area, where the literature includes examination of effects on panel statements on practice within the United States and internationally. Some of these evaluations are first-rate sources of information.

Other sources were written reports by persons directly involved in the panels (for example, Stalford, 1987 on the NIE Convening Process experience) and interviews (for example, with Dr. Houser, the organizer of the National Stroke Association convening process and Dr. Alan Kraut, organizer of the extensive American Psychological Society research agenda consensus effort. To reduce errors in listening and recording, a copy of the relevant sections of this report was sent to the interviewees for corrections, changes, and additions. Appendix II lists the interviewees and contacts and should help identify what or who may have been missed in the consensus panel network.

For reasons of time, only the experience of domestic agencies is considered, even though the Departments of Defense and State may have strategies for using consensus panels that would be useful to OERI. A final exclusion is that most of the experience will be fairly recent: this is not a comprehensive history of consensus panels.

TYPES OF AVAILABLE INFORMATION ABOUT CONSENSUS PANELS

PUBLISHED

1. Descriptive: Materials describing the purpose of the panel, its mission, the general process. This type of information is most frequently available. However, published information even of a descriptive nature is lacking for some extremely interesting panels. Further, there are many gaps even in the descriptive literature, such as the basis for panelist selection and the time and costs required for panels.

2. Reflective: Reports of experiences with the panel, the problems encountered, how they have been solved, observations on the results of panel recommendations. This type of information is less frequently available. There are, however, some exceptions where federal project officers responsible for consensus panels have written in depth about their experiences and published the reports or presented them in professional meetings. Some of these examples can be considered high quality self-evaluations.

3. Analytic: Third-party assessments of the reliability, validity, and cost/effectiveness of the panel. These analyses are rare. The National Institutes of Health medical practice panels have been most extensively scrutinized by third parties including members of the evaluation community such as Lee Sechrest and Paul Wortman. However, according to Shekelle, there is no systematic research on such a basic question as the key underlying assumption that a different panel drawn from the same pool of selected panelists held at the same time period would yield similar results. Data on the US Department of Education Program Effectiveness Panels also have been analyzed by third parties.

4. Research: There is quite large body of research on individual versus group processes of decision-making in the psychological, sociological, administrative literatures that are relevant to consensus panels. In some analyses of consensus panels, some of this research is analyzed and integrated into conclusions.

UNPUBLISHED

Interviews have proven useful in obtaining descriptive information and reflective information. They do not seem to uncover much analytic data, either as unpublished third-party reviews or unpublished internal data. The information obtained this way is reported, but noted as unverified.

APPENDIX II: AGENCY, ORGANIZATION, INDIVIDUAL CONTACTS OUTSIDE OF OERI, ED

Many colleagues contributed by suggesting other contacts, sending materials, and providing new information through telephone interviews. My deepest thanks to each and every one!

Agency for Health Care Policy and Research, HHS. Executive Office Center 2101 East Jefferson Str., Suite 501, Rockville, MD. 20852, 301-227-8364. Dr. Sloat (301-594-1447) is an institutional memory on grants awarded re consensus processes who described the Shekelle award. Contact Dr. Carole Hudgings (301-594-4015) for information on AHCPR consensus development guidelines.

American Psychological Society, APS. Dr. Alan Kraut, Executive Director. APS is developing a national research agenda in the social and behavioral sciences through a consensus conference process. Detailed information on impact of "The Human Capital Initiative" agenda. TEL: 202-783-2077; FAX: 202-783-2083.

Anderson, Dr. Richard. Director, Center for the Study of Reading. Expert on reading. Led OERI-funded work on "Becoming a Nation of Readers." FAX: 217-244-4502.

Calamus, consulting group, of Judy and Steve Selig, in Niederland, Colo. Uses consensus groups for the resolution of scientific/technical issues. TEL: 303-258-7888.

California Associations of Nonprofits (Alan Fox), PO Box 1478, Santa Cruz, CA, 95061. TEL: 408-458-1955, FAX 408-458-9486. Called him to find out if he knew of any reports on use of consensus panels by nonprofits for research priority setting or dissemination. (He did not know of any.)

Environmental Protection Agency (Tom Kelly, Director for Policy and Evaluation; TEL: 202-260-7500; FAX 202-260-0513). Source of within EPA Consensus-Panel names. See also Lynn Luder, formerly of OTA (202-260-6995) with experience in OTA extensive use of consensus panels as well as EPA examples. EPA main number: 202-260-2090/Office of Public Information, 260-2080/Office of Pesticide, TEL: 703-305-5805.

> Dr. Kelly described two EPA uses: regulatory negotiations (draft regulations) and a science court (policy dialogues). Panels represent diverse affected groups--those whose ox may be gored, and they come under a written contract to see consensus, accepting ahead of time the procedure of a facilitated committee, and binding themselves to accept/sign a product to be published. Key person: CHRIS KIRTZ, 202-260-7565. Kelly's address is Office of Regulatory Management and Evaluation, Stop 2131, U.S. Environmental Protection Agency, 4th and M St. SW, Washington, DC, 20460.

Dr. Lynn Luderer Desautels, Director of the Risk Communication Project, OPPE, Stop 2131, US Environmental Protection Agency, 401 M St., SW, Washington, D.C. 20460 (TEL: 202-260-6995; FAX: 202-260-0513) gave extensive leads to Chris Kirtz and the negotiated rulemaking experience (TEL: 202-260-7576) and Deb Dalton (TEL: 202-260-5495). She also referred me to OTA/Peter Blair; the Keystone Center in Colorado; Calamus in Niederland, Colo; World Wildlife Fund dispute resolution/resolve. These are listed separately.

Food and Drug Administration (Mike Taylor, Director for Policy, Program Planning and Evaluation, TEL: 301-443-2854; FAX: 301-443-5930/ contact, Jennie). See also Donald Sauer, Director ACM, TEL: 301-443-3370, FAX: 301-443-5161/contact, Sharon Holsten for linkage to FDA groups using consensus panels. FDA general: TEL: 301-443-1130 and 301-443-1544. See FAXED list of seven FDA staff contacted, none of whom had much to say. FDA seems to be is writhing under the self-inflicted wound of a NAS report commissioned by the outgoing FDA Director which is said to be highly critical of FDA's use of expert panels, and staff were not inclined to discuss their views.

HIRI (Dr. Thomas Backer, TEL: 818-501-5432/FAX: 818-501-4638) sent references and list of foundation contacts. Excellent source of information on the psychodynamics of groups and decision-processes, and classic work of Dr. Edward Glaser.

Jefferson Center for New Democratic Processes (Amy Richards, Administrative Assistant). 364 Century Plaza, 1111 Third Avenue S., Minneapolis, Minn. 55404-1007. TEL: 612-333-5300, FAX: 344-1766. Detailed information on reports of the 60 plus citizen juries already held and on the processes; also sent video of a jury and detailed time/cost information.

Keystone Center, Colorado. Uses mediation and dispute resolution through consensus panels. TEL: 303-468-5822.

National Institute on Disability and Rehabilitation Research. Dr. Ellen Blasiotti (TEL: 202-205-9800/FAX: 202-205-8515; U.S. Department of Education, 400 Maryland Avenue, SW, Washington, DC 20202) Source of excellent, detailed information on NIDRR procedure, time, and costs.

National Dissemination Association. Dr. Max McConkey, Executive Director, 4732 North Oracle Road, Suite 217, Tucson, AZ 85705 (608-888-2838). He arranged for an "all hands" announcement in the NDA newsletter and sent a copy of their annual program to see if members have research/experiences with consensus panels. One response received.

National Institutes of Health. Dr. Charles Sherman, Tel. 301-496-1143, Fax 1-301-402-0420 is a source of evaluations on the NIH consensus process. He also sent extensive public information materials on the panels which are also available via 1-800-NIH-OMAR.

National Stroke Association, (Karyl Newman/Gary Houser; 303-771-1700; FAX 303-771-1886.) Information on time, costs, procedures of the consensus panel on emergency treatment for stroke.

Office of Technology Assessment. Uses scientific/technical panels to answer congressional questions. Consensus group. See Peter Blair, who leads energy work, for experiences with panels. 1-202-228-6260.

Reid, Ethna R. Dr. Reid is Director of the Exemplary Center for Reading Instruction and responded to my quest for consensus panels published in the NDA Update. Together with her husband, she has written a foresighted article in Reader, on the use of consensus processes for moving from research to practice in education. Her article includes intriguing consensus-building exercises, based on the NIH panels.

Shekelle, Paul. Senior Researcher at RAND, Santa Monica. Extensive research on necessary and appropriate clinical practice consensus statements. TEL. 310-393-0411, FAX: 310-393-4818.

Szczypkowski, Ronald B. President, MAGI Education Services Inc. Rye Plaza, 601 Midland Avenue, Rye, NY 10580. TEL 914-921-1969 FAX 914-921-4347. Extensive experience with New York State Sharing Success System, state validation panels which have many consensus panel features, including diverse stakeholder participation and emphasis on evidence of results. Generously initiated an independent search via phone to NIH, FDA and EPA, reporting that while these agencies said they used public advisory committees, all indicated they did not use consensus panels or ever heard of any like term.

World Wildlife Found, operates CEDR, a mediation, dispute resolution group using consensus panels to develop scientific statements. 1-202-861-8330. Key person is Gail Bingham, Vice President for the group, and its Newsletter, RESOLVE.

APPENDIX III: REFERENCES

Acknowledgement: Two extremely valuable lists were prepared by David Tsuneishi of the Reference Services Branch, Research Library Division, Library Programs, Office of Educational Research and Improvement, U.S. Department of Education, using the term, "consensus development". The first search checked educational sources; the second, all other sources.

Agency for Health Care Policy and Research. AHCPR Grant Announcement. Washington, D.C.: Author, March 1992.

Allender, Jerome S. "Educational research: A personal and social process." Review of Educational Research. Vol. 56 (2), Summer 1986, pp. 173-93.

American Psychological Society. "The Human Capital Initiative," Special Issue of the Observer, February 1992.

American College of Chest Physicians/Society of Critical Care Medicine. 'Ethical and moral guidelines for the initiation, continuation, and withdrawal of intensive care." Chest, Vol. 97(4), pp. 949-958.

American Psychological Society. "The Human Capital Initiative: Vitality for Life--- Psychological Research for Productive Aging." Special Issue of the Observer, December 1993.

American Diabetes Association. Consensus Development Conference on Diet and Exercise in Non-insulin Dependent Diabetes Mellitus. Diabetes Care, Vol. 10(5) Sept/Oct/ 1987, pp. 639-644.

American Psychological Society. "The Human Capital Initiative: The Changing Nature of Work," Special issue of the Observer, October 1993.

Asch, S. M. and C. E. Lowe. "The consensus development program." Knowledge: Creation, Diffusion, utilization. March, 1984, pp. 369-385.

Atkinson, R.C. and G. B. Jackson (eds.) Research and Education Reform: Roles for the Office of Educational Research and Improvement. Washington, D.C.: National Academy of Sciences, National Research Council. National Academy Press, 1992.

Backer, Thomas E. and Kirk B. O'Hara. Knowledge Transfer and Utilization in Health Care: A Review for the National Emergency Medical Services for Children Program. Los Angeles, CA; Human Interaction Research Institute. December 1989.

Baker, M.R. "When does consensus exist in expert panels?" J. Public Health Medicine, Vol. 13(4), Nov. 1991, pp. 343-344.

Bankart, C.P. and S.W. Powers, "Individual decisions and group consensus," Journal of Social Psychology, Vol. 126, 1986, pp. 369-374.

Berry, John and Judy Quinn. "Consensus out of chaos." Library Journal, Vol. 116 (13), August 1991, pp. 42-46.

Bland, D. J. "A comparison of two evaluation processes: a convening activity versus the traditional." Paper presented at the 1982 meeting of the Evaluation Research Society, Baltimore.

Brennan, T.A. "Practice guidelines and malpractice litigation: collision or cohesion?" Journal of Health Politics, Policy and Law, 16 (1), Spring 1991, pp. 67-85.

Brook, R. H., M.R. Chassin, A. Fink, et al. "A method for the detailed assessment of the appropriateness of medical technologies," International Journal of Technology Assessment in Health Care, Vol. 2, 1986, pp 53-63.

Brown, D. Piecing Together the Evidence: How an NIH "Jury" Reaches a Verdict on Medical Advice. The Washington Post, Health Section. May 19, 1992, pp.12-15.

Chassin, M., "How do we decide whether an investigation or procedure is appropriate?" in Hopkins, A. (Ed.) Appropriate Investigation and Treatment in Clinical Practice, Royal College of Physicians, London, 1989.

Chassin, M.R., R. E. Park, A. Fink, et al., "Indications for selected medical and surgical procedures---a literature review and ratings of appropriateness: coronary artery bypass surgery," Santa Monica, CA.: RAND, R-3204/2,CWF/HF/HCFA/PMT/RWJ. 1986.

Christofeel, K. K., P.C. Scheidt, P.F. Agran, J. F. Braus, E. McLoughlin, J.A. Pauslon. Standard definition for childhood injury research: excerpts of a conference report. Pediatrics, Vol. 89 (6, Part 1) June 1992, pp. 1027-1034.

Clarfield, A. M. and Foley, J. M. "The American and Canadian consensus conferences on dementia: is there consensus?" Journal of the American Geriatric Society. Vol. 41 (8), August 1993, pp. 883-886.

Cleeman, J. I. and Lenfant, C. New guidelines for the treatment of high blood cholesterol in adults from the National Cholesterol Education Program: From controversy to consensus. Circulation, Vol. 76(4), Oct. 1987, pp. 960-962.

Clinton, J. Jarrett, McCormick, Kathleen and Besterman, Jacqueline. "Enhancing Clinical Practice: The Role of Practice Guidelines." American Psychologist, Vol. 49 (1), 1994, pp 30-33.

Cooper, Harris and Hedges, Larry V. The Handbook of Research Synthesis. New York: Russell Sage, 1993.

Crosby, Ned. The Basis of the Citizens Jury Process. Minneapolis, Minn.: Jefferson Center for New Democratic Processes. July 7, 1993.

David, J. H., Kameda, T., C. Parks, et al., "Some social mechanics of groups: the distribution of opinion, polling sequence, and implications for consensus," Journal of Personality and Social Psychology, Vol. 57, 1989, pp. 1000-1012.

Davis, J.H., R.S. Tindale, D.H. Nagao et al., "Order effects in multiple decisions by groups: a demonstration with mock juries and trial procedures," Journal of Personality and Social Psychology, Vol. 47, 1984, pp. 1003-1012.

Demarest, Elizabeth J. and Sheryl E. Stein. A National Agenda for Education Research and Development: Why Not An A+. Unpublished draft paper, April 1993.

Department of Education. Exemplary Educational Programs and Practices Funded by the National Diffusion Network. Washington, D.C.: Author, 1994.

Demarest, Elizabeth J. and Sheryl E. Stein, "A National Agenda for Educational Research and Development: Why Not an A- Plus?" Office of Educational Research and Improvement, U.S. Department of Education, February, 1994 Revised Draft.

Department of Education. "What is PEP Validation?" Washington, D.C.: Author. Undated.

Department of Health and Human Services, Office of the Inspector General. The NIH Consensus Development Program: Dissemination of Findings through Medical School Continuing Education Activities. Dana Miller, Project Leader, Timothy Corbett, Lead Analyst Boston Regional Office, Tel. 617-565-1050. February 1994, OEI-01-91-01760.

Donnellon, Anne et al. "Communication, meaning and organized action." Administrative Science Quarterly. Vol. 31 (1), March 1986, pp. 43-55.

Durand-Zaleski, I., F. Bonnet, H. Rochant, P. Bierling, F. Lemaire. "Usefulness of consensus conferences: the case of albumin." Lancet, December 5, 1992. 340(8832), pp. 1388-1390.

Dyson, Freeman J. "Science in trouble", American Scholar, Vol. 62, No. 4, Autumn 1993, 497-512.

Emihovich, Catherine et al. The making of a model: Utilizing consensus in formative evaluation. Paper presented at the American Evaluation Association Annual Meeting, Boston, MA: October 14-17, 1987.

Environmental Protection Agency. Negotiated Rulemaking at the Environmental Protection Agency. Washington, D.C.: Author, November 1, 1993.

Environmental Protection Agency. Discussion on the Use of Consultation and Consensus-Building Processes for Implementing the Clean Air Act of 1990. Washington, D.C.: Author, March 4, 1992.

Environmental Protection Agency. Fact Sheet on Negotiated Rulemaking/Regulatory Negotiation. Undated.

Environmental Protection Agency. U.S. Environmental Protection Agency Advisory Committee Charters, Rosters and Accomplishments. Washington, D.C.: Author, June 1992.

Environmental Protection Agency. An Assessment of EPA's Negotiated Rulemaking Activities. Washington, D.C.: Office of Policy, Planning and Evaluation, Environmental Protection Agency, December 1987.

Ernst, Nancy D. NIH Consensus Development Conference on Lowering Blood Cholesterol to Prevent Heart Disease: Implications for Dieticians. Journal of the American Dietetic Association, Vol. 85 (1), May 1985, pp. 586-588.

Fadale, LaVerna and Gene M. Winter. Evaluating developmental studies programs in two-year colleges: a consensus approach. Paper presented at the American Vocational Association Annual Meeting, Las Vegas, NV, December 6-10, 1987.

Feinstein, A. R. "Fraud, distortion, delusion and consensus: the problems of human and natural deception in epidemiologic science." American Journal of Medicine. Vol. 84 (3, Part I), March 1988, pp. 475-478.

Fern, K., Jr. Increasing understanding of public problems and policies. Paper Presented

at the 39th National Public Policy education conference, New Orleans, LA; September 18-21, 1989. (Oak Brook, Ill: Farm Foundation)

Food and Drug Administration. List of Public Advisory Committees. Undated.

Forgione, Pascal D., Jr., Martin E. Orland. A Guide to Improving the National Education Data System: Executive Summary. Washington, D.C.: National Center for Educational Statistics. March 1991.

Franz, M.J. "Consensus statement on Diabetes Control." Journal of the American Diabetic Association. Vol. 92 (9), Sept. 1992, pp. 1136-1139.

Fraser, G. M., D. Pipel, S. Hollis, J. Kosecoff, and R. H. Brook. "Indications for cholecystectomy: the results of a consensus panel approach." Quality Assurance Health Care (ENGLAND), Vol. 5 (1), March 1993, pp. 75-80. (Israel)

Glaser, E. M. "Using behavioral science strategies for defining the state-of-the-art," Journal of Applied Behavioral Science, Vol. 16, 1980, pp. 79-92.

Green, Alan. Communicating in the 80s: new options for the nonprofit community. Washington, DC: Benton Foundation, 1983.

Hargens, Lowell L and Hagstrom, Warren O. 'Scientific consensus and academic status attainment patterns." Sociology of Education. Vol. 55 (4), Oct. 1982, pp. 183-196.

Hargens, Lowell L. "Further evidence on field differences in consensus from the NSF peer review studies." American Sociological Review, Vol. 53 (1), Feb. 1988, pp. 157-160.

Hartman, C. The limits of consensus building in the land use planning process. In j. I. DeNeufville, The Land Use Policy Debate in the United States, New York: Plenum Press, 1981, pp. 205-207.

Havelock, Donald G. The consensus development conferences of the National Institutes of Health as a model for education and other fields. In Ward S. Mason (Ed.) Proceedings of a Conference on Disseminating Findings of Research on Postsecondary Governance and Finance. November 21, 1986. College Park, MD: National Center for Postsecondary Governance and Finance.

Havercamp, M.J., D. J. Torell, and W.P. Evans. Using PGA for consensus building in public land disputes. Journal of Soil and Water Conservation, Vol. 48 (2), Mar/April 1993, pp. 90-93.

Hedges, L. V. "Research synthesis: the state of the art", International Journal of Aging and Human Development,' Vol. 19(2), 1984, p. 85-93.

Hillman, B. J. "The consensus of committees," Invest. Radiology, Vol. 27 (1), Jan 1992, p. 2-28.

Italian Psycho-Oncology Society (SIPO) Consensus Development Conference: Assessment of the quality of life in cancer clinical trials. Tumori, vol. 73 (3), June 30, 1992, pp. 151-154.

Jacoby, I. "Biomedical technology: information dissemination and the NIH consensus development process." Knowledge; Creation, Diffusion, utilization." Vol. 5. 1983, pp. 245-261.

Jacoby, I. "The consensus development program of the National Institute of Health: current practices and historical perspectives." International Journal of Technology Assessment in health Care. Vol. 1 (2), 1985, pp. 420-432.

Jefferson Center for New Democratic Processes. America's Tough Choices: Citizens Jury Report on Federal Budget Panel. Minneapolis, Minn.: Jefferson Center for New Democratic Processes. 1993.

Jennings, B. Possibilities of consensus: toward democratic moral discourse. Journal of Medical Philosophy., Vol. 16 (4) August 1991, 447-463.

Kanouse, David E. et al., Changing Medical Practice Through Technology Assessment: An Evaluation of the NIH Consensus Development Program. Santa Monica, CA: The RAND Corporation, March 1989.

Khaleel, Ibrahim Adams. Significant issues confronting vocational education: implications for UCVE institutions. University of Missouri: University Council for Vocational Education, Nov. 1990.

Klein, Susan S. "Research and Practice: Implications for Knowledge Synthesis in Education." Knowledge: Creation, Utilization, Diffusion. 11 (1) 1989, 58-78.

Klein, Susan S., Norman Gold and Charles Stalford. "The Convening Process: A New Technique for Applying Knowledge to Practice." Educational Evaluation and Policy Analysis. (2), 1986, 189-205.

Knodel, J., W. Sittitrai, and T. Brown, "Focus group discussions for social science research: a practical guide with an emphasis on the topic of aging," University of

Michigan: Population Studies Center, No. 90-3, 1990.

Kosecoff, J., D. E. Kanouse, W. H. Rogers, L. McCloskey, C. M. Winslow, R. H. Brook. "Effects of the National Institutes of Health Consensus Development Program on physician practice." Journal of the American Medical Association, Vol. 258 (19) pp 2708-2713.

Kruger, W. C. Building consensus for rangeland uses. Rangelands, Vol. 14(1), Feb. 1992, pp. 38-41.

Kruger, R. A. Focus Groups: A Practical Guide for Applied Research. Newbury Park, CA: Sage, 1988.

Larson, J.L. and E. M. Rogers. "Consensus development conference: a cross-cultural study," Knowledge: Creation, Diffusion, Utilization. 1984. 537-548.

Laughlin, P. R. and R. P. McGlynn, "Collective induction: mutual group and individual influence by exchange of hypotheses and evidence," Journal of Experimental Social Psychology, vol. 22, 1986, pp. 567-589.

Leape, L., R. E. Park, J. P. Kahan, et al., "Group judgments of appropriateness: the effect of panel composition," Quality Assurance in Health Care, 1992.

Lomas, J., G. M. Anderson, K. Domnick-Pierre, E. Vayda, M.W. Enkin, W. J. Hannah. "Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians?" New England Journal of Medicine, Vol. 321 (19), Nov. 9, 1989, pp.1306-1311.

Lomas, J. Words without action? The production, dissemination, and impact of consensus recommendations. Annual Review of Public Health, 1991, Vol. 12, pp. 41-65.

Lomas, J., G. Anderson, M. Enkin, et al. "The role of evidence in the consensus process. Results from the Canadian consensus exercise." Journal of the American Medical Association, Vol. 259, 1988, pp. 3001-3005.

Long, Steve. Vocational Implication of UTI among People with SCI. Paper prepared for the National Institute on Disability and Rehabilitation Research, undated.

Lundberg, Carig C. and Alan M. Glassman. "The informant panel: a retrospective methodology for guiding organizational change." Group and Organization Studies. Vol 8 (2), June 1983, pp. 249-264.

Lynch, Richard L. et al. Using the nominal group technique to determine priorities for vocational education research. Blacksburg, VA: Virginia Polytechnic Institute and State University. August 1987.

Martorella, Peter H. "Consensus building among social educators: a Delphi study." Theory and Research in Social Education. Vol. 19(1), 1991, pp. 83-94.

McKinney, Kay. Improving Math and Science Teaching: A report on the Secretary's October 1992 Conference on Improving Mathematics and Science Teaching and Instructional Resources. Washington, D.C.: U.S. Department of Education, February 1993.

Meyes, R.A. "Testing persuasive argument theory's predictor model: alternative interactional accounts of group argument and influence," Communication Monographs, Vol. 56, 1989, pp. 112-132.

Miner, William M. and Dale E. Hathaway. World Agricultural Trade: Building A Consensus. Halifax, NS: Institute or Research on Public Policy and Institute for International Economics., 1988.

Montgomeries, John Z. Treatment of Urinary Tract Infections in Patients with Spinal Cord Injury. Paper prepared for the National Institute on Disability and Rehabilitation Research, Nov. 1991.

Moskowitz, Howard R. "The case for and against recruiting external panels for developmental product evaluation," Journal of Food Quality. Vol. 2(3), September 1979, pp. 231-234.

Mullan, F. and I. Jacoby. "The town meeting for technology: the maturation of consensus conferences." Journal of the American Medical Association. Vol. 253 (August), 1985, pp. 1068-1072.

Munoz, Ricardo, Hollon, Steven D., McGrath, Ellen, Rehm, Lynn P., and VandenBos, Gary R. "On the AHCPR Depression in Primary Care Guidelines: Further Considerations for Practitioners." American Psychologist, Vol 49 (1), 1994, 42-61.

National Institutes of Health. Consensus Development Conference Statement: Differential Diagnosis of Dementing Diseases. Washington, D.C.: National Institutes of Health, Vol. 6 (11), July 1987.

National Institutes of Health. Consensus Development Conference Statement: Total Hip Joint Replacement. Washington, D.C.: National Institutes of Health, Vol. 4 (4), 1982.

National Institutes of Health. Consensus Development Conference Statement: Cochlear Implants. Washington, D.C.: National Institutes of Health, Vol. 7 (2), May 1988.

National Stroke Association. "NSA Leads Emergency Treatment of Stroke with Consensus Statement," Be Stroke Smart, Vol. 10, No. 2, 1993, p.1-2.

National Institute on Disability and Rehabilitation Research. Protocols for Choosing Low Vision Devices: January 25-27, 1993 Consensus Statement. Vol 1, No. 4. Washington, D.C.: Author. Undated.

National Institute on Disability and Rehabilitation Research. Supported Employment for People with Severe Mental Rehabilitation: May 23-26, 1993. Consensus Validation Statement. Washington, D.C.: Author. Undated

National Institute on Disability and Rehabilitation Research. Strategies to Secure and Maintain Employment for People with Long-term Mental Illness. Consensus Statement, September 21-23, 1992. Vol. 1, No. 3. Washington, D.C.:Author. Undated.

National Institute on Disability and Rehabilitation Research. Augmentative and Alternative Communication Intervention: National institute on Disability and Research Consensus Statement. Washington, D.C.; Author. 1992.

National Academy of Sciences, National Research Council, National Institute of Medicine. Consensus Development at the NIH: Improving the Program. Washington, D.C.: National Academy Press, 1990.

National Institute on Disability and Rehabilitation Research. Augmentative and Alternative Communication Intervention: March 23-26, 1992. Consensus Statement. Vol. 1, No. 2, Washington, D.C.: Author. Undated.

National Institutes of Health. Panel and Conference Chair Responsibilities. Undated.

National Institutes of Health. Guidelines for the Planning and Management of Consensus Conferences. Undated.

National Institutes of Health/OMAR. Checklist for the Consensus Development Conference Program Coordinator. Undated.

National Institutes of Health. Guidance for Using the NIH Consensus Program Clearinghouse, List of Past Conferences, Statement Order Form, and 1993-1994 Conference Schedule. Undated.

National Institutes of Health. NIH Consensus Statement: Early Identification of Hearing Impairment in Infants and Young Children. Washington, D.C.: Author, Vol. 11 (1), March 1993.

National Institutes of Health. Consensus Statement: Triglyceride, High Density Lipoprotein, and Coronary Heart Disease. Washington, D.C.: National Institutes of Health, Vol. 10 (2), February 1992.

National Institutes of Health. Consensus Development Conference Statement: Estrogen Use and Postmenopausal Women. Washington, D.C.: National Institutes of Health, September 1979.

Nolen, S.L. and Dimmick, W.F. EPA's control technology approach to assisting states and regions with air toxic problems: five case studies. U.S. Environmental Protection Agency: Air and Energy Engineering Research Laboratory, Research Triangle Park, NC. Toxicological Indicators of Health, Vol. 6 (5), October 1990, pp. 257-267.

Okun, M.A. and Stock, W.A. "The relationship between health and life satisfaction: a comparative analysis of three research synthesis techniques." International Journal of Aging and Human Development, Vol 19(2), 1984, pp. 81-83.

Perry, S. and Wilkinson, S.L. "The technology assessment and practice guidelines forum: A modified group judgment method." International Journal Technology Assessment and Health Care. Vol. 8(2), Spring 1992, pp. 289-300.

Poster, John B. The limits of consensus. Paper presented at the Eastern Educational Association annual meeting, Boston, MA, February 25-March 1, 1987.

Press, S. J."Qualitative controlled feedback for forming group judgments and making decisions," Journal of the American Statistical Association, 1978.

Ralph, John and M. Christine Dwyer. Making the Case: Evidence of Program Effectiveness in Schools and Classrooms. PIP 89-819. Washington, D.C. Department of Education, 1989.

Roth, Christopher and Seth Chaiklin. Emerging consensus in novice physics problem-solving research. Paper presented at the American Educational Association annual meeting, Washington, D.C.: April 20-24, 1987.

Schulberg, Herbert C. and Rush, A. John. "Clinical Practice Guidelines for Managing Major Depression in Primary Care Practice: Implications for Psychologists," American

Psychologist, Vol. 49 (1), 1994, 34-41.

Shekelle, Paul G. Assessing Appropriateness of Expert Panels" How Reliable? Grant HS07185, U.S. Department of Health and Human Services, Agency for Health Care Policy. Santa Monica, CA: RAND Corporation, 1992.

Sherman, C. R., A.L. Potosky, K. A. Weis, and J. H. Ferguson. "The consensus development program: detecting changes in medical practice following a consensus conference on the treatment of prostate cancer." International Journal Technology of Assessing Health Care. Vol. 8 (4) Fall 1992, pp. 683-93.

Stalford, Charles B. "Congruence in theory and practice: the convening process and the consensus development program," Knowledge: Creation, Diffusion, Utilization. Vol. 9(1), September 1987, pp. 4-19.

Stalford, Charles B. "Congruence in Theory and Practice," Knowledge: Creation, Diffusion, Utilization. 9 (1), 1987, 4-18.

Stein, Marcy et al. The Beginning Reading Instruction Study, Washington, D.C.: U.S. Department of Education, June 1993.

Stone, James R. Conference consensus evaluation: National marketing education curriculum conference. Atlanta, Georgia, December 1984.

Suphin, H. Dean and Camp, William G. "A model for building consensus on the applications of microcomputers in agricultural education." Journal of Vocational Education Research., Vol. 15 (3), 1990, pp. 65-79.

Tomlinson, Tommy. Hard Work and High Expectations: Motivating Students to Learn. Washington, D.C.: U.S. Department of Education, April 1992.

Torell, D. and Havercamp, M.J. Consensus Building: A Guide for Public Land Facilitators. University of Nevada, Reno: Nevada Cooperative Extension, 1991 (91-14).

Toseland, Ronald W. et al. "An evaluation of decision-making methods in task groups." Social Work, Vol. 29(4), Jul-Aug, 1984, pp. 339-46.

U.S. Department of Health and Human Services. Secretary's Panel on Teen Pregnancy Prevention. Washington, D.C.: Author. 1989.

U.S. General Accounting Office. Vocational Rehabilitation: Evidence for Federal Program's Effectiveness is Mixed. Washington, D.C.: Author. GAO/PEMD 93-19, 1993.

U.S. Department of Education. Improving America's Schools Act of 1993. Washington, D.C.: Author. Sept. 13, 1993.

Vinokur, A., et al., "Group decision making by experts: Field study of panels evaluating medical technologies." Journal of Personality and Social Psychology, 1982, 49, 70-84.

Ward, Barbara. The National Assessment Approach to Objectives and Exercise Development. Denver, CO: Education Commission of the States, Sept. 1980.

Ward, Spencer A. and Linda J. Reed. Knowledge Structure and Use: Implications for Synthesis and Interpretation. Philadelphia, PA: Temple University Press, 1983.

Weingarten, S. and A.G. Ellrodt. "The case for intensive dissemination: adoption of practice guidelines in the coronary care unit." QRB Quality Review Bulletin, Vol. 18(12), December 1992, pp. 449-455.

Woolf, S. H. Practice guidelines: a new reality in medicine. II: Methods of developing guidelines. Archives of Internal Medicine, Vol. 152 (5), pp. 946-952.

Wortman, P., A. Vinokur, and L. Sechrest, "Do Consensus Conferences Work? A Process Evaluation of the NIH Consensus Development Program. Journal of Health Politics, Policy and Law, Vl. 13 (3), Fall 1988, pp. 469-498.

Woteki, Catherine E. "Building consensus for nutrition education research," Journal of Nutrition Education, vol. 11(1), Jan/March. 1979, p. 7.

APPENDIX IV: CONSENSUS BUILDING AND CONSENSUS PANELS – APPLICATIONS FOR RESEARCH AGENDA SETTING

The National Institute of Education Cognitive Skills Research Agenda: A Consensus Building Approach in Reviewing the Field

Background: In 1978, a decision was made at the level of the Office of the Director, then Dr. Patricia Graham, to place high priority on cognitive skills and to issue a grants announcement that would be appropriate for a five to ten year period of inquiry: enough time and resources, it was hoped, to make notable strides in the field. Cognitive skills was given this priority in large part because in national assessment after assessment, students were doing far better on rote items than on items requiring cognitive skill, be it in reading comprehension, mathematics problem solving, or in areas such as planning and problem solving generally.

The Process: The process matched the vision. A steering committee was selected from leading researchers in this field, "balanced" primarily to reflect different schools of thought and research traditions such as ethnographic and experimental. About six months went into preparing a framework intended to synthesize the state of knowledge in the field, the logic being that priorities would emerge from analysis of the gaps. An extensive list was prepared of potential paper-writers, commissioned to prepare first-rate journal-quality syntheses on each topic of the framework. This list was balanced to achieve gender and racial diversity within the pool of first-rate researchers in the field. About a year went into commissioning the papers, preparing drafts, reviewing drafts by the steering committee and preparing for an international conference on cognitive skills. The conference was to serve as a broad consulting or consensus development activity intended to assure that (a) the state-of-knowledge was thoroughly and accurately presented; (b) the research opportunities identified would be endorsed by key researchers in this field or disputes known well in advance of the grants announcement, and (c) the groundwork would be laid for receipt of exceptionally high calibre, on-target proposals. About 200 persons attended the five day conference, and through paper-sessions, seminars, discussion groups, a broad consensus did emerge. During the next year, the steering committee prepared the papers for book publication and the grants officer prepared the announcement, one intended to stand for about five to ten years.

Cost: The total cost was about two years of staff time at the GS 12-15 level, and about $500,000 all total in contract funds for all aspects of the planning task: roughly about $700,000 in current dollars.

Lessons Learned: What happened next was that a wealth of excellent proposals were received but less than anticipated were funded by the agency due to a change in leadership and loss and reallocation of funds. Within the Education Department, only the group leading the steering committee (the Learning Research and Development Center at

the University of Pittsburgh) incorporated much of the agenda in its center-funded research. However, the agenda was largely realized because staff dispersed to other agencies such as the Department of Defense, taking the ideas (and the mailing list of researchers) with them. About ten years later, the pendulum swung back again and a "what has been learned about cognitive skills" conference was held, using the 1980 report as the benchmark. Much had been learned, but on populations other than elementary and secondary school children whose needs had prompted the initiative and whose performance on reasoning, comprehension and problem-solving continued to be problematic.

The American Psychological Society's Research Agenda Committee:
A Case Instance of a National Research Consensus Conference

Background: In 1992, the American Psychological Society (APS) published "The Human Capital Initiative" identifying six priority areas for research in the behavioral social sciences that had been endorsed by over 70 national associations, representing thousands of researchers. In 1993, APS brought out the first two of the projected series of detailed research agendas that would take the issues in each of the priority areas from topic to research study. The scope of the effort, and its intended long-term impact---which includes dissemination of eventual results---may be unprecedented as an instance of using the consensus process to shape a national research agenda.

According to Dr. Alan Kraut, Executive Director of the American Psychological Society, the APS initiative was stimulated by a dud. The National Academy of Society published in 1988 the long-awaited product of the deliberation of a panel of experts, "Research Opportunities in the Behavioral Sciences." It sparked no interest, no one was listening: as an effort to stimulate researchers, to help agencies set priorities, to interest Congressional committees or other funding sources, it failed.

Thinking about the need for such a document, APS concluded that among the problems with the report were that no priorities were set, too wide a scope was taken, and ideas ran too vast a gamut. Truly, "if everything is important, then nothing is."

APS decided to go for an actual consensus on an inspiring and tightly prioritized research agenda, one requiring tradeoffs and resolution of conflicts among diverse groups. Further, APS wanted the process to build so strong a consensus among so many different stakeholder groups, that attention would be paid by virtue of the quality of the report and the quantity of interest groups involved.

Overall Process: APS invited over 70 research organizations to participate (at their own expense) in the first Behavioral Research Summit held in January 1990 at Tucson, Arizona. At this meeting, the groups "...unanimously endorsed the development of a national research agenda that would help policy makers in federal and other agencies set funding priorities for psychological and related sciences." (APS, 1992, p. 7)

The leaders of the major organizations were asked by the 70 participants to appoint a Steering Committee to draft a report expressing the national purpose articulated. The Steering Committee reported at the second Behavioral Science Summit held in January 1991 in Houston, Texas. In February 1992, the consensus report, "The Human Capital Initiative", was published. It focuses, as the title indicates, on human capital development as the top national priority for the behavioral sciences. Six areas of national concern are identified: (1) productivity in the workplace, (2) schooling and literacy, (3) the aging society, (4) drug and alcohol abuse, (5) health, and (6) violence in America.

According to Dr. Kraut, this report has had lots of visibility in academia and among Appropriation Committee members. For example, Senator Mikulski, whose committee has oversight of the National Science Foundation, had some nasty things to say in 1993 about NSF---the only nice thing was her enthusiasm for NSF support of the APS national consensus on research. Fully 20% of the NSF Behavioral and Social Science Research increase in dollars went to the Human Capital Research priorities.

As another example, the head of the relevant Senate authorization committee wrote the Director of the National Institute on Aging a letter enthusiastically commending NIA support of the effort, and asking the Director to look at the report and let the Senate know how they would respond. The incoming Director of NIA has meet with APS, noting how useful the report is not only for specific research priorities but also for educating him in new areas.

The second stage of the effort is well-underway. Using the consensus process, and the overall Human Capital Initiative vision, detailed long-term research agendas are being prepared through APS for agencies including the National Institute of Drug Abuse, the National Institute on Aging, and the National Science Foundation Behavioral and Social Sciences unit.

The 70 organizations participating in the agenda development have, in the meantime, remained solidly behind it. Dr. Kraut noted that ED/OERI is aware of the initiative, and has called in the context of planning research priorities for the Laboratory and Center competition.

Nitty-Gritty Process: One aspect of nitty gritty is money. According to Dr. Kraut, this has been a shoe-string operation, with direct costs of about $40,000. I estimate additional costs to include perhaps $40,000 more in APS staff time and about $70,000 in time, travel and other expenses donated by the representatives of the 70 organizations over the two-year period. The National Institute of Health direct cost grant clearly levered almost 300% in "matching" funds. Other institutes contributed smaller grants. The direct costs went mostly $200 at a time in helping organizations with scanty budgets afford travel to the summit; for steering committee travel; and for printing many copies of the widely distributed "Human Capital Initiative" report.

Another aspect is the psychodynamics of groups and procedures. Dr. Kraut observed that each summit has been a learning experience. A particularly challenging problem has been weaning participants from the ever-flowing milk of simply adding research issues to the agenda to the solid food of setting a very few priorities...in other words, forcing confrontation, decisions, negotiations rather than caving in and adding topics to placate everyone. To achieve this, APS used several strategies. One was selecting as leaders researchers of great stature, superb negotiating skills, and modest egos. Another was saturating participants with a litany from federal research administrators: "The train is

running. You will tell me the direction where you want it to go or I will tell you where it is going. Priorities must be set."

Yet another nitty-gritty question is format and style. From the beginning, APS wanted an attention-grabber that could be understood without an interpreter by non-researchers on Congressional staff and in the administration. They got it: the agenda is readably laid-out, well-illustrated, and it talks informatively about national concerns such as schooling and literacy, and the aging society. Further, the title is catchy and it had the credibility of unanimous support from the 70 plus groups listed prominently on p. 5 (the title page) with pictures of the participants at each Summit meeting facing the title page. The not-too-subliminal message was "We are behind this document and we will be looking at what you do next."

The readable way in which quite complex research is brought to bear may be illustrated in Chapter 3, Schooling and Literacy, which includes a section on "The Psychology of the Individual: Behavior, Mind and Brain":

> "While wishing for a 'smart pill' may be a fool's desire, we have in fact learned that some mental processing depends on certain neurotransmitters in the brain that are modifiable, sometimes as simply as changing how we eat...Expanding this knowledge could enable us to overcome obstacles to learning... This is exciting work, occurring under a variety of labels, and deserves to be pushed ahead." (1992, p. 17)

And another issue: the link between researchers and practice communities. In the second phase, now underway, quite tight links are being built in. Development of the specific research agenda for aging included representatives from organizations such as the American Association of Retired Persons as well as federal agency and congressional staff as observers and reviewers.

Dr. Kraut summarizes the 8 step process:

1. Select a broadly recognized person, with those superb negotiating skills and willing to volunteer plenty of time, to chair (or co-chair) the entire enterprise.

2. Bring together 6 to 8 steering committee members who write down their views of research priorities. The chair organizes this; the draft is circulated and revised until consensus is reached.

3. Invite representatives of the widest possible set of organizations with some concerns for research on the issue (in this case, the issue of the aging society) to a 1 to 1 1/2 day meeting.

4. They read the draft document in advance. Their task is to revise and rewrite it...which they do, often pulling all nighters in an effort to achieve consensus.

5. About 3 weeks later, after the steering committee edits, the revised draft is circulated to all groups

6. Following comments, suggestions, concerns, the chair completes a final edit.

7. The final draft is circulated for approval or not

8. In Dr. Kraut's experience, at this point, almost all organizations approve.

In between, particularly at steps 2, 4 and 6, the Chair uses the phone, fax, and e-mail, conference calls and other approaches to achieve consensus through negotiation within the guidelines of adherence to the national vision and the need to establish tight priorities.

Lessons Learned: This process is still evolving. Dr. Kraut is disappointed that private foundations haven't reacted as enthusiastically as Congressional committees and is exploring why. He also cautions that the initial impact---which seems close to what APS intended---may not hold up for all second step research agendas. And although the initial work was done on almost less than a shoe-string, agencies will need to share with APS, whose budget is extremely limited, the costs of developing detailed agendas in all six of the broad national concerns.

Further, systematic "compared to what" information is unavailable. The APS agenda may have notable overlap with existing agenda derived in more traditional ways in the various agencies and research organizations or it may be considerably different. The research agenda resulting from the APS process may be so puissant in getting funds to do research, however, that the question of what a different process might have yielded is moot. Being a bit untactful, if the process gets the myriad special interest organizations to work together, getting behind a single agenda, instead of going separately for their own slice of the pie (or the crumbs), whatever may be lost in scientific cutting edge issues may be regained in the long-run through additional support.

Next Steps for APS: From OERI/ED's perspective, APS's newest initiative may be of considerable interest. In June 1994, APS will hold a major national conference on meta-analysis and research synthesis, building on the 8 year effort sponsored by the Russell Sage Foundation which has involved such experts as Dr. Fred Mosteller, Dr. David Cordray and Dr. Richard Light. OERI has supported many research syntheses and research-into-practice efforts, that seem---from a sample sent to me---not to have used the formal meta-analytic techniques (perhaps for excellent reasons associated with the technical or substantive aspects of the specific topics). The June 1994 APS conference would give OERI almost a free ride on the expensive, extensive Russell Sage work on the

leading edge in this area. (Dr. Kraut noted he had been called to discuss these issues and opportunities with someone in OERI but that no follow-up had been made after the initial conversation.)

APS notes that the research agenda is only the first step in "what is intended to become a continuing process of bringing systematic research to bear on problems of national interest" (1992, p. 1). Meta-analysis may speak both to research agenda development, and to dissemination.

APPENDIX V: CONSENSUS BUILDING AND CONSENSUS PANELS AND DISSEMINATION OF RESEARCH KNOWLEDGE TO IMPROVE PRACTICE

The Program Effectiveness Panel

Background: The Program Effectiveness Panel is something of a hybrid. It is consensus-panel-like in having a group of experts systematically reviewing evidence of effectiveness for an educational project, product or practice; it is also consensus-panel-like in that a decision is reached. However, it is not consensus-panel-like in several respects. First, there is no systematic review of all relevant evidence about the topic, only a review of information about a specific application. Second, a majority vote carries the decision and the panel is not enjoined to come to maximum possible consensus. Third, at present, the panel does not meet in person and, except for mail/telephone communication to clarify points, there is no discussion.

The panel is included in this review because it represents an early approach to a federal look at evidence to reach a decision about-educational innovations through a group process that in the initial years was interactive. Also, it has influenced some states such as New York to develop much more consensus-panel-like and elaborated processes of their own (Szczypkowski, 1994; New York State Education Department, undated). In New York state, the Sharing Success Program which has been fully operational since 1980, has become the flagship mechanism for identifying exemplary programs and practices and disseminating them throughout the state (Doino-Ingersoll, 1993).

Process: Begun around 1973 in an effort to place some floor of reviewed evidence of effectiveness under U.S. Office of Education enthusiasm for an educational approach, the Panel originally was staffed primarily by highly trained evaluators in the Department, the panel meeting in a sort of science court form, and the review focused on whether claims of effectiveness were supported by the evidence. Although unanimity was not required, an effort was made to resolve points of disagreement, particularly on what the evidence was and its strengths, before voting. Department project officers were responsible for deciding if the approach was innovative (or ho-hum); for determining that the innovation existed in practice and not just on paper; and for initial review of evidence of effectiveness to be sure enough was there to be worth the Panel's time. Any topic could be considered from any Department of Education program. Evidence was to be presented in 10 pages, read ahead of time by the panel. Panels met in something of a hearing format, usually with the Program Director, Program Evaluator and Project Officer as witnesses to answer Panel questions. A simple majority vote determined whether the Panel concluded evidence was strong enough to be convincing with regard to effectiveness. Designation as an exemplary program won entry into the Department's

National Diffusion Network, and the opportunity for the developer to apply for modest funds to help make the program available to interested educators.

At present, the Panels are more diverse, particularly in involving practitioners and non-Federal reviewers as well as evaluators and Federal panelists. Criteria are more diverse, including but extending beyond evaluative evidence of program effectiveness. The evaluative evidence itself reflects changes in the field of evaluation, particularly with regard to qualitative information. Similarly, a wider array of innovations may be submitted, and federal funding is not a requirement. Panel members read the materials and send in their recommendations rather than meeting. Designation as an exemplary program continues to offer recognition, funds and other opportunities for developers to help others adopt their innovation, in addition to information about the approach entering the Technical Assistance Centers networks.

Further, an organization, the National Dissemination Association, has been established to help match adopters and adoptees, and carry out the process of adapting the innovation to its new setting.

Selection by the Program Effectiveness Panel may, however, get considerably less fanfare than selection by the Department's other strategies for recognizing and encouraging outstanding education, such as the Teacher of the Year awards and selection as an exemplary school through the Blue Ribbon School Recognition Program. Processes for these selections are recommendations from elementary and secondary school leadership to a state review group and each state is entitled to designate a certain number of teachers and schools. While effectiveness is certainly prominent in the intent and language of the school recognition award, the evidence is not primarily empirical, or judged with great rigor. Much is based on a consensus of beliefs about best practice.

Costs and Time: In its present form, the Program Effectiveness Panel involves only modest direct costs (mailing the packages of applications to reviewers) and indirect costs (staff time for reviewing packages, mailing them out, collating the results). A rough estimate would be about $2,000 per review. While proposal preparation takes time, the review itself would take about 6 to 8 weeks.

Lessons Learned: The PEP and NDN both have been fairly extensively studied, including research and evaluations. The research has aimed primarily at the panel process, such as the panelist consistency and the characteristics of more and less successful proposals. Results from these studies have been incorporated into the current approaches. For example, there is greater clarity about the acceptability of qualitative as well as quantitative designs, measures and analyses and greater emphasis on the innovativeness and quality of the project or product being reviewed. Studies of the longer-term impact of the program have identified indirect benefits such as greater attention to evidence of effectiveness and evaluation quality and direct benefits such as

attention to selected programs and their adoption or adaptation by other schools...using, rather than reinventing, the educational wheels. Little is known systematically about the replicability of effectiveness in the adoption sites, although the program may have some rich natural variation data.

For the New York State program, an evaluation of the 1-5 year longevity and impact of program adoptions showed that about 75% of the replication sites had implemented their selected program; about 90% of these continued to use the program in whole or part; and over 80% reported they were satisfied or very satisfied with process aspects of the replication. Impact on outcomes such as student learning or teacher skills were not studied (Doino-Ingersoll, 1993).

In terms of PEP's value for a possible ED/OERI consensus panel, methodological developments such as the evaluation criteria and conceptualization of the distinctions among promising and exemplary products and practices could be useful in the future.

CITIZENS' JURIES: A CONSENSUS PANEL PROCESS REGARDING BROAD POLICY ISSUES

Purpose: State and national issues with federal, state and local implications have been the focus of an experimental program supported since 1974 by the Jefferson Center for New Democratic Processes of Minneapolis, Minnesota. The originator of the Citizen's Jury, Ned Crosby, is concerned with giving more ordinary citizens a voice in determining "how we should live together" that can be heard above the special interest clamor. Such pressures can affect more formal hearing processes such as those used (albeit with over 500 experts and thousands of other witnesses) during the almost year-long effort on health care reform that led to the President's recommendations.

Process: The jury process has been evolving. For example, there is now an oversight committee of former jurors to help maintain the integrity of the process. To get "significant audiences" to pay attention to the jury, public relations techniques appropriate to the event are used. In the 1992 Pennsylvania election jury, for example, the process was structured to get Senator Arlen Spector to participate. Once an issue has been selected (by the Center leaders), exactly how the questions get framed have varied from jurors, surveys, sponsoring groups, to Center staff. How the information gets organized has varied from juror control to staff provided frameworks: all involved, however, witnesses. Jurors usually select witnesses from a short list prepared by the "case managers" or center staff. The jurors themselves are selected first at random, and then from this list, to represent a "microcosm" of the relevant community; jury sizes have ranged from 7 to 24. The process is evaluated by the jurors themselves after each project, and through the analyses of Center leadership (Crosby, 1993).

According to the Center, media attention to the panels have included articles in the *Washington Post, New York Times, Los Angeles Times, USA Today, Newsday, the Philadelphia Inquirer, Minneapolis Star-Tribune, St. Paul Pioneer Press, and Des Moines Register.* In addition, ABC-TV and the National Public Radio, C-Span, CBS Radio and CBC radio have covered the panels. For the October 1993 health care jury, a broader public could listen to a one-hour PBS presentation.

The first national jury, held in January 1993 to review the federal budget, may be a useful specific example. This Citizens' Jury involved selection of 24 jurors, chosen at random but balanced to reflect the characteristics of the national electorate with regard to ethnicity, age, gender, income, geographic location, education, political affiliation and attitude toward taxes and spending. The jurors met for two days of discussion, both with moderators and separately. On the third day, formal presentations of major competing approaches were made. On the fourth day, the jury broke into smaller panels to discuss special issues, and on the fifth day, the jurors voted on changes they wanted in the competing Republican and Democratic approaches. The findings of the panel were released at a news conference on the fifth day and presented in written form to Congress

and the Administration (Jefferson Center, 1993).

Costs: The Center estimates total costs for a 5-day, 24-person national jury at almost $225,000. Direct costs involve preparation and administration; payment of honoraria (about $100 per day) and expenses for the panelists; and payment for the expert witnesses and moderators, travel, honoraria, living expenses, preparation and publication times. Dissemination costs were probably relatively modest for the tax and budget jury; for the health care jury, the cost of videotaping and editing for a PBS presentation would have to be added ($115,000).

The Center itself was founded by Ned Crosby, whose family helped found General Mills; sponsors and underwriters of the Federal Budget Panel included Patricia Benn and Ned Crosby, the Jefferson Center, the law firm of Opperman, Heins and Paquin, Jill Buckley and Associates Public Affairs and Himel Horner Public Relations.

Time: Based on the reports in addition to the five days of meetings, preparation required about six months and follow-up, about four months in elapsed time.

Lessons Learned: This is rather on the borderline as an example of a consensus panel dealing with research for dissemination purposes. First, consensus is not required as it would be in a formal jury although the intensity of discussions mean individual judgments are informed far beyond anything that is likely to emerge from an opinion poll, a speculation consistent with the published statements by jurors. Further, although the aim is citizen representativeness, the size (24 people) is too small, in a statistical sense, to permit national projections of what informed views would be like among various groups, had they the time and inclination to spend five days on a topic such as the federal budget. While the information provided included policy analysis and applied research, basic research is not prominent. And lastly, while the long-term goals of the Center are significant, the stakes for a specific panel seem to be low: there are no consequences of the jury's decision other than another voice---however well-informed or representative---in complex debates.

The processes used still may be illuminating for the Department of Education such as the random selection of jurors, the time allotted to the panel meeting, and the sequencing of how information is presented and decisions reached and the issues raised in Cosby's extremely thoughtful analysis of the enterprise. At the least, this two decades long experiment suggests the applicability of the jury process to social issues, and a notable respect for the ability of "ordinary" people to weigh complex information and discern the good, the true, and the beautiful.

THE NATIONAL INSTITUTES OF EDUCATION CONVENING PROCESS: FROM CRAFT KNOWLEDGE/RESEARCH TO PRACTICE

Background: In 1981, the Superintendent of the District of Columbia Public Schools, Floretta Dukes McKenzie, approached the National Institute of Education for help in solving a problem (Klein, Gold and Stalford, 1986). The School Board required an evaluation of a competency based approach (the student progress plan) that had looked good on paper but experienced notable problems in practice, both in administration and teaching. The approach taken by the Institute, with full support of the Superintendent, as an experiment in evaluation was to create a convening process.

Process: This process began with identification of nine administrators from other school districts with similar populations who had hands-on, extensive, authoritative experience with competency based programs--some successful and some not successful. These consultants were to reach a group consensus. Extensive pre-planning with District representatives, the panel, and NIE staff lead to a 2 l/2 day site visit to the District. Consultants read documents, interviewed stakeholders at all levels, and met together. They were ready, at the end of their visit, for an oral report with 15 recommendations, followed by a draft and then the final report.

In 1983, the convening process was used again in response to a request from the Mississippi State Department of Education for help in planning how to carry out their 1982 comprehensive education reform legislation. In this instance, the convening process focused on school accreditation and teacher certification. State Department of Education officials were the primary stakeholders and the consensus panel selected from their peers, this time with both researchers and experts with hands-on practical knowledge--the peers of the Mississippi educators.

At the end of a 3 day site visit, prior to which the panel had read about the convening process, the Mississippi legislation, and related pre-planning papers, the panel gave their recommendations, followed by the written reports.

Results: In the case of the District, most of the panel's recommendations were adopted. The instance is particularly intriguing since a concurrent, independent and "traditional", survey-based evaluation had also been commissioned. Some but not all of their recommendations were consistent with those of the Panel. In the Mississippi case, panel recommendations were more general. Almost all of panel recommendations were acted upon in recognizable ways, according to the evaluation of the panel process and its results by an OERI staff member.

Time and Costs: These panels had both direct and indirect costs. Direct costs included payment to the panelists, and their travel and expenses. At about $5,000 per panelist, a ten member panel would cost about $50,000 in today's money. The indirect costs to the LEA and the SEA include the extensive preplanning involving many stakeholders, participation in the approximately 3 days of interviews and site visits, and follow-up

meetings and reviews of documents. A rough estimate would be that 20-30 persons were involved at each site, for approximately a week each. If the average cost per week of relatively high-level time runs $1,000 (loaded with indirect costs such as sick and annual leave), about $20,000 to $30,000 was invested by each agency in staff time. Time for the Institute included design and carrying out of the approach. Assuming that about half of the time went into designing the convening process and half to carrying it out, and that about 4 staff members were involved to varying degrees, about 24 weeks total of staff time was required, at about $1,000 (loaded) or roughly $24,000 for the implementation of the process.

Elapsed time for the actual convening process is relatively short, compared to more traditional forms of either evaluation or knowledge synthesis. Selecting and recruiting the consultants would require about 3 weeks, preparing the materials and schedule another 4 weeks, the convening event about a week, and follow-up, about 10 weeks of elapsed time---about 4 to 5 months.

Lessons Learned: Expertise in the convening process was largely vested in one staff member who left the Institute; enthusiasm for a collegial approach in working with LEAs and SEAs, and the quick-turnaround resources to do it diminished. The US Department of Education has not initiated similar efforts recently.

However, as a process, several lessons learned have been identified as essential:

- identification of practical problems that can be solved with currently available knowledge
- selection of colleagues with experience with successful and unsuccessful analogs in highly similar situations
- people who would form a "natural system of support"
- use of a formal, structured process to achieve consensus which includes planning, orientation, on-site information gathering and analysis,
- summary analysis and
- follow-up

Putting this together, the process is relatively fast, but it is not cheap. Further, relatively little has been written about the specific processes used to achieve consensus, and the ease or difficulty with which this was obtained. Some degree of challenge is suggested by the emphasis in the written reports on orientation and on assuring that all involved understand and agree to the objectives of the convening process. The process does appear to be effective, however, judged by the report impact, by consumer satisfaction, and by the reported integration of craft knowledge and research achieved in the discussions.

THE PRINCETON CITIZENS' COMMITTEE: A CONSENSUS PANEL GRAPPLES WITH SCIENCE AND PUBLIC POLICY

Background: A remarkable qualitative report (Dyson, 1993) describes the consensus panel commissioned by the Mayor of Princeton, New Jersey to deal with citizen distrust of biologists' management of recombinant DNA. The rules established in 1976 regarding what experiments could take place under various levels of containment did not satisfy the citizens and thus the local political authorities in Princeton or in one other community---- Cambridge, Massachusetts. In both places, authorities appointed citizens' committees "...to study the hazards and to advise the authorities as to whether and under what conditions recombinant DNA experiments should be permitted in the town." (p.518)

Process: The Princeton committee involved eleven people, chosen for representativeness and diversity with regard to gender, ethnicity, education and employment. The committee worked intensively for four months, hearing testimony and "educating ourselves about practical details of disease control, bacterial epidemiology, laboratory design, and experimental protocol." Meetings were public, and actively attended.

Results: A consensus was not achieved. Eight panelists wrote a majority report recommending experiments go ahead; three recommended experiments requiring a specially built laboratory (due to extreme danger) be prohibited. Because a consensus was not achieved, the town council examined the issues themselves in detail, and nine months after the citizens' committee reports were delivered, voted by five to one to follow the majority report. However, the citizens' committee did succeed in providing a neutral ground for debate between university biologists and the citizens and kept the dispute out of the arenas of confrontations and the courts, and set a tone for trying to reach a compromise "based on mutual respect."

Costs: Information is not available; apparently the citizens were not reimbursed for their work.

Time: To the four months elapsed time, about a month seems to have been required before and after the committee concluded their formal work.

Lessons Learned: As seen by the author, three big lessons were learned. One was the importance of listening and particularly the value of person-to-person contacts and the give-and-take of open discussion. The second was that "sincere and well-informed people may have fundamentally divergent views about the ethics of science." (p. 520) The primary reason for the minority report was not so much the technology of safety, risk and containment as larger ethical, moral, and philosophical issues that were imminent in the circumscribed charge to the citizens' committee and which the minority strongly felt could not be ignored, such as whether insertion of genes from higher organisms into lower organisms in itself constitutes a dangerous breach of evolutionary barriers. The third lesson, as seen by Dyson, is that "..the good and evil faces of science should be openly acknowledged", recognizing the inherent risks, unpredictabilities and uncertainties of conclusions based on science, or about science.

THE NATIONAL STROKE ASSOCIATION'S CONSENSUS PANEL RESEARCH INTO PRACTICE AS FAST AS POSSIBLE!

The leaders of the National Stroke Association believed that stroke should be treated as an emergency, and that too often, stroke patients were given stabilizing treatment, rather than immediate therapeutic treatment and rehabilitation: "The sooner we begin to treat the brain after stroke, the better our chances to minimize unnecessary secondary damage." (1993,. p. 1). To give this conclusion greater prominence, the National Stroke Association created a consensus panel of ten top national stroke authorities. The panel's task was to develop a consensus statement on early treatment, one which could be endorsed by leading medical organizations. The specific topic was how soon emergency evaluation and treatment should start; the consensus that the weight of evidence showed that intensive treatment (as appropriate to the case, including pharmacological treatment) within the first six hours after a stroke can notably limit brain damage.

Process: NSA leadership selected 8 distinguished national experts as their panel, chaired by the President of the leading neurological association. The panel met on four occasions, primarily to flesh out the issues and assign specific areas of writing to each. The panel shared and revised the manuscript four times, using a modified Delphi process. The final revision was sent to large medical societies for cosponsorship and endorsements, as well as comments. The final manuscript was reviewed again by these societies. It was released in a press conference at the annual meeting of the American Academy of Neurology; all the panelists attended as did many members of the invited media such as the *New York Times* and the Associated Press. A video news release had been prepared and was widely shown. Copies of the consensus statement were sent to 50,000 cardiologists and 110,000 emergency room physicians, EMTs and nurses. Public awareness kits, videos and brochures had been prepared; 10,000 of these had been sent out together with a coordinated mailing to American Association of Retired Persons health coordinators, Area Agencies on Aging and other health care providers.

Results: The statement has been endorsed by the American Academy of Neurology, the American Association of Neurological Surgeons, the American Society of Neuroimagiing, the Congress of Neurological Surgeons, the International Stroke Society, the National Institute of Neurological Diseases and Stroke, and the World Congress of Neurology. Six months after the release, the statement has been adopted as a standard for practice in over 40 hospitals. Feedback has been extensive and unanimously positive.

Costs: The Association received a $40,500 grant to cover editorial salaries and panelist travel. My estimate is that indirect costs in additional staff time may have run another $40,000. In addition, the Association obtained $245,000 for dissemination. These funds probably leveraged another $200,000 in assistance in dissemination.

Time: Start time in this instance was from appointment of the 8 member planning group to the final manuscript: from December 1992 to May 1993.

Lessons Learned: This was the first time the National Stroke Association had used a consensus process as a tool for dissemination; they would do it again. Having absolutely top people is seen as crucial to achieving the visibility, credibility and impact the Association feels the effort has had, because authority carries great weight in the health community. The effort had the full support of the Director of the National Institute of Neurological Disorders and Stroke, Dr. Murray Goldstein, who encouraged the Association to go for it, because the information was too crucial to wait for the four years they believed it would take using the NIH/OMAR panels. (Houser, personal communication, 1993).

THE NATIONAL INSTITUTE ON DISABILITY AND REHABILITATION RESEARCH: A CONSENSUS VALIDATION PANEL FOR RESEARCH INTO PRACTICE

Background: The Department has at least one instance of a consensus panel approach to dissemination that is close to, and in fact has been adapted from, the NIH/OMAR model. In January 1992, the National Institute on Disability and Rehabilitation Research held the first conference of a three year program "...to evaluate and synthesize available scientific information and improve the dissemination of findings from research....[the aim is] to help close the gap between research and practice, and to encourage the release of the results of research at an earlier point than has been the case traditionally." (NIDRR, undated, p. 1). NIDRR's term for the process is a consensus validation conference and the statements are called consensus validation statements---a nuance that is accurate in comparison to the NIH/OMAR nomenclature.

The history of the program was NIDRR concern that there was not much to tell or report from the work of the research centers supported by the Institute. There were project reports, yes, but nothing that said something major had been accomplished, nothing to give NIDRR name recognition as a place that got something done to solve problems, nothing to attract media attention.

Several consensus conference statements already have been published. The topics include: prevention and management of urinary tract infections among people with spinal cord injuries (NIDRR, 1992); augmentative and alternative communication intervention (NIDRR, 1992); strategies to secure and maintain employment for people with long-term mental illness (NIDRR, 1992); protocols for choosing low vision devices (NIDRR, 1993); and supported employment for people with severe mental retardation (NIDRR, 1993).

As described by NIDRR,

> The Consensus Validation Statements...are prepared by a non-advocate, non-Federal panel of experts, based on (1) resource papers prepared preliminarily by experts; (2) testimony presented by researchers, clinicians, and consumers during a one-day public hearing; and (3) a day of closed deliberations by the panel, during which the consensus statement is prepared. This statement is an independent report of the panel and is not a policy statement of NIDRR or the Federal Government." (NIDRR, undated, p. 1)

More Detail on the Process: Ms. Ellen Blasiotti of the NIDRR Dissemination and Utilization Program provided further details.

- Topics are selected by the NIDRR Director, choosing from among suggestions by project officers on what is timely and where there is enough information to make a statement. The Director names a planning committee to refine the topics and frame the questions for the proceedings.

- The committee develops a framework around the issue, usually five or six questions such as "What are the best standard clinical and functional assessment practices in vision rehabilitation for adults with low vision?" and "What assessment and referral services do primary care specialists provide for adults with low vision?"

- The committee then crafts workscopes for commissioned background papers for each question (and any other paper needed); recommends paper-writers; and recommends ten people to be on the panel.

- The chair of the planning committee is always the chair of the conference panel itself.

Panel members always include some people with disabilities, people from vocational rehabilitation agencies, and others who have credibility and an appropriate background but who are not paper-writers.

- Before the conference, the panel members review the papers. The panel chair and the conference contractor ask panel members to take a crack at answering one or two of the questions and to come to the panel meeting with suggestions for further research. This saves times in drafting a final statement, by giving panel members a starting point rather than having to write from scratch. NIDRR believes that if the panel had to start from scratch writing a statement, it might take weeks. The pre-drafted answers are seen as serving as a basis for discussion, (put into the computer ahead of time) but they are not regarded by NIDRR as a pre-drafted consensus.

- The conference works as follows: the ten member panel meets for dinner the night before to get their charge from the Director and discuss the process. The first day and a half, they hear pubic testimony and discuss papers and may question and otherwise interact with witnesses and each other. That night, they go into executive session, with secretarial help from the contractor, and write or revise the draft consensus statement as they see fit. The new material and changes are loaded into the computer, and the full statement printed out for the full presentation.

- o This consensus validation statement is not final. It is presented by the panel in a for-the-media session. The panel listens to changes suggested by the audience. The members then go back into executive session and revise the statement if necessary. The final draft is given to the contractor for typing. It then goes to the Education Department for clearance.

Ellen Blasiotti noted that the Department clearances hold up the release a lot: many picky changes. It takes about three months from the end of the editing process to release. However, the topics are not time sensitive, so this usually doesn't affect the accuracy of the contents.

Costs: NIDRR uses Conwal, Inc. as a contractor to handle all aspects of the conference. The cost is about $100,000 for their work (including up to about $1,000 for the paper writers and their travel; travel for panelists; secretarial services; etc.) My estimate is that if NIDRR staff time is added, including pro-rated time for process development and topic selection, plus time for "negotiating" clearance and GPO publication, the costs of each conference would run about $115,000 upwards. Ms. Blasiotti noted that NIH has many more resources, and each Institute and center within each Institute has the opportunity to put on an OMAR consensus conference. Further, NIH has the money to pay for time and travel to get many more witnesses, much more testimony. The NIH Centers also do the professional staffing so a conference contractor is needed primarily for conference arrangements with much less substantive interaction with paper-writers and they put a lot more money into dissemination, such as videotapes.

Time: When Department clearance time is considered, about 12-18 months seems about the length from when proposals are prepared to when the consensus validation statement is released. Ms. Blasiotti estimates that elapsed time when the planning committee first meets to statement release is about 6 months.

Results: The program has not been systematically evaluated. Ms. Blasiotti observed, however, that "...there are no real consequences" in the sense that NIDRR can impose no sanctions and there are no liabilities for not paying attention to state-of-the-art knowledge. The consensus validation statements are seen by NIDRR as only guidelines, and the demands of consumers (those served by rehabilitation services) may be the most feasible route available to NIDRR for implementation of the guidelines. As the consumers get better informed, they may demand service based on the consensus validation statements.

That is, from NIDRR's perspective, "the panels are only suggesting new knowledge, and do not seek to change practice, only to inform it." Ms. Blasiotti commented that even the NIH/OMAR panels do not have 100% impact, although they carry a lot of weight.

NIDRR tries to get media attention by having the consensus validation panel members use their contacts with professional newsletter editors and with the rehabilitation community, but the statement releases are not big media draws. NIDRR has found that it is hard to get media attention in Washington for disability topics by agencies of less stature than NIH.

Lessons Learned: Within the Education Department context, a pilot study for OERI actually is available. The most promising features may be that it is possible to select solution-oriented, laser-focussed topics where a state-of-knowledge statement can be useful. For example, the value of the assisted communication devices is a readily recognizable hot topic. Among the concerns, however, are (1) the extent to which research compared to craft knowledge most influences the statements [the statements I saw made no reference to any specific research. The one commissioned paper I read on the vocational implications of urinary tract infections among people with spinal cord injury concluded there was relevant research on UTI among people with SCI but little relevant research on the vocational implications of UTI for persons with SCI]; (2) the extent to which selection of panelists "facilitated" consensus relative to reaching an agreement among researchers who are kindly disposed and those more skeptical; (3) the extent to which the quite long list of needed research studies gives a message of "we don't really know a lot about this" too strongly coupled with a message of "adopt this practice"; and (4) the disparity between three months of clearance by the Department and a product that is an independent statement by the scientific/practitioner community that is regarded as free of federal policy considerations.

This latter issue balances the desire for sponsorship (as a Department and NIDDR activity from which NIDRR can get well-deserved credit for moving the field along) with the varying degrees of changes that Department clearances can imply. NIDRR believes it has reached an ideal compromise and that little actually is tampered with in clearance. However, the NIH process that brings the agency due credit seems to go to great lengths to assure consensus panel independence, integrity and trustworthiness.

In neither case is there identification for the reader of which research (if any) relied upon by the consensus panel was primarily sponsored by the agency (NIDRR or NIH). Arguably if garnering recognition for supporting a major body of breakthrough research was a primary goal, clearer links between agency funding for the research and the breakthrough would make the case better than anonymity. Thus the credit and recognition involved would seem to adhere primarily on serving as a convening agency bringing knowledge to bear on a signally important issue according to the considered judgment of a most credible, impartial, authoritative group, operating under scientific restraints only.

THE NATIONAL INSTITUTES OF HEALTH CONSENSUS CONFERENCES FROM RESEARCH TO MEDICAL PRACTICE

Medical practice often follows a fairly structured sequence from problem, such as treatment of high blood pressure, through research, to experimental protocols on animals, to clinical trials on volunteers, to wide-spread practice. At each stage, there usually is rigorous scrutiny of the evidence, including the replication of results by several independent researchers.

Some bodies of evidence proceed relatively smoothly, with little dispute. In other areas of practice, the controversies may be more salient and the stakes considerably higher. One element of these stakes is that federal medical payments are not provided for experimental treatments. The size of this stake is huge, including the Public Health System facilities, Indian Health Services, Medicare, Medicaid and the Department of Veterans' Affairs. Further, many health insurers follow suit, and do not authorize payment for treatments the federal payment system considers experimental.

The consensus program was initiated to resolve controversial issues in clinical practice using quality scientific data, a review appropriate given the importance of medical care to life itself, and the total national costs of such care. National Institutes of Health began the consensus panel reviews in 1977. The NIH term for the overall effort is the Consensus Development Program (CDP); for a specific review, a Consensus Development Conference (CDC).

In the words of the oversight committee of the 1989 Rand evaluation of the NIH Consensus Development Program, which give a sense historically of CDC's purpose:

> CDC's avowed purpose is to publicly evaluate scientific information concerning biomedical technologies and arrive at consensus statements that will be useful to health care providers and the public at large, and that will serve as contributions to scientific thinking about the technologies under consideration.
>
> CDP is more than an assessment program. It is also a communication program to the professional community and the public. It aims to disseminate the results of assessment to health care professionals (as well as researchers) throughout the country in order to improve the state of professional practice." (Kanouse et al., 1989, p. v)

How CDP Currently Works: CDP continually evolves and is fine-tuned as experience develops, additional research findings on the consensus process come in, and recommendations of major evaluations commissioned by NIH are considered. As CDP currently works,

- NIH has a component called OMAR (Office of Medical Applications and Research) within the Office of the Director. OMAR has staff resources and money to fund consensus development conferences, the know-how to make them work well, and widely promotes information on CDCs to journals and the media, a "know how" that can foster though not guarantee coverage.

- The Directors of the various Institutes are charged with knowledge creation on important topics. The Directors and their staff can propose to OMAR topics for a consensus conference. The Directors have many incentives to do so. The process is an effective way to get knowledge into practice; the culture of NIH is results oriented and CDP is an established part of the use-of-knowledge-to-improve-medical-practice strategies.

- When the OMAR and the Directors agree that the topic is highly important AND a body of evidence seems ripe for major assessments, OMAR puts the consensus conference on the conference schedule, with the actual meeting usually about 1 to 2 years off. A crucial element in the decision is that some element of controversy exists.

  OMAR staff emphasize that there are three main criteria: the topic must be of public health importance, sufficient data must exist for a CDC panel review, and there must be controversy. Because of OERI interest in topic selection, the OMAR (NIH, undated) criteria as printed are noted. A topic should meet the following selection criteria:

  > It should have public health importance; it should affect or broadly apply to a significant number of people.
  >
  > Controversy or unresolved issues should surround biomedical/scientific aspects of the topic that would be clarified by the consensus approach, or there should be a gap between current knowledge and practice that a CDC might help to narrow.
  >
  > It must have an adequately defined and available base of scientific information from which to answer the conference questions and to resolve the controversies insofar as possible.

It should be amenable to clarification on technical grounds and the outcome should not depend mainly on the subjective judgment of panelists.

Additional positive considerations are health care, cost impact, preventive impact, and public interest (NIH, pp. 4-5)

- OMAR allocates funds for the conference, and assigns an OMAR liaison to help the appropriate Institute scientists understand how the consensus development panel process works. OMAR has put together what seems like an excellent checklist for Institute staff, as well as its "Guidelines for the Planning and Management of NIH Consensus Development Conferences."

- It is crucial that the coordinator from the Institute be a senior staff member with great knowledgeability in the area of science under consideration. The Institute must have scientists of stature and knowledge who know where issues are in the field, what the research strengths and limits are, and the landscape of the research in terms of the controversies and who is involved. In other words, OMAR provides consensus conference management and dissemination expertise; the Institute must have senior staff members who are themselves extremely knowledgeable scientists in the topic.

- The sponsoring Institute's Information Office also assigns a staff person to work with the OMAR Director of Communications. Conference publicity, media coverage and planning for dissemination of the conference statement begin right from the start of consensus development planning.

- The coordinators identify other health-related Federal organizations that might co-sponsor the conference. These are asked to join the planning committee. A panel chairperson is carefully selected, and non-Federal employees knowledgeable about the scientific/technical issues involved are invited to join the planning committee. The planning committee has four main duties: developing the consensus questions, developing the conference agenda, nominating speakers, and nominating panelists.

- The panel chairperson is very carefully selected. Ideally, the chairperson

  "should be a knowledgeable and prestigious figure in the field of medical science under consideration, but should not be identified with a strong advocacy of the conference topic or with research that might be presented to answer any of the conference questions." (NIH, undated, p.6)

- The conference witness or speakers are selected for scientific expertise on the topic and "care must be exercised to include the presentation of opposing data." Abstracts of the speakers' presentations are due at least two months before the conference.

- The panel is selected for a range of expertise on the topic and is expected to include research professionals; health professionals such as practicing physicians; methodologists such as epidemiologists; and public representatives such as patients. They must be free from conflict of interest financially or by prior organizational advocacy or promotional positions or with any research likely to be cited in the testimony so they are not judging their own work or that of anyone else where conflict might be raised, such as one of their staff members. Federal employees can not be panel members to avoid any appearance of undue federal involvement. Panelists must be U.S. citizens.

- Usually, panel size runs between 9 and 16 members with 12 to 14 as an optimum working group.

- The planning committee works to refine the key questions, which often deal with efficacy, risks, clinical applications and directions for future research. They are framed so that answers can come from scientific information, not subjective judgments or opinions and it will be evident that consensus has been achieved.

- The conference itself includes a preliminary panel orientation; a 1 1/2 day plenary session with speakers, witnesses, and open discussions among panelists and the audience. On the night of the first day, the panel meets in executive session to begin drafting the consensus statement. In the afternoon of the second day, the panel goes into executive session and finishes drafting the statement.

- The statement is presented publicly the morning of the following day, modified if the panel sees fit on the basis of audience comments, and then formally adopted. A news conference is held to disseminate the panel's findings to the media. A news release is also distributed at this time. The panel and chairperson are available for media interviews and press conference follow-up.

- The written form of the statement is proofed by OMAR for style, syntax and clarity; sent to all panelists for final comments; sent to the panel chairperson and is then considered final. This is a fast forward process and does not involve any clearances.

- Extensive advanced planning goes into dissemination including systematic identification of target audiences, identification of their key newsletters and major professional journals, encouragement of interested parties to attend the conference and present testimony, and getting ready for fast mailout of the statement to key organizations and publications. Occasionally, dissemination may involve developing a video of the proceedings, depending on the topic and funding.

- The NIH Office of Education can award continuing medical education credit for members of the audience. This is done routinely for every CDC.

- OMAR is responsible for conference evaluation. This includes planning evaluation protocols to get baseline/pre-information before the conference and at intervals afterward (an interrupted time series design.)

Costs: OMAR staff state that the cost per conference is about $150,000 to $160,000 on the average. The actual logistics such as arranging travel are handled through a competitively-award standing contract for support of the consensus development panels.

Based on a description of the process, costs per conference should include payment of travel and honoraria for the panels, any overhead-type costs for the logistics contract, commissioned papers, staff time to develop materials and organize panel meetings and proofing/printing of the statement. The price might go up if one were to add the direct and indirect costs of the processes of proposal review and development within each Institute from which the final slate for each year is selected and the post-statement dissemination efforts which may take some staff time such as writing articles relevant to the CDP or specific panel statements for various journals or presentations at medical societies. Also, the OMAR "overhead" in developing the methodology, conducting evaluations and such probably should be amortized over an appropriate period (5 years?).

Each year, the Institutes pay a set amount to OMAR's budget. This supports the CDP and is not tied to a specific consensus development conference planned for the year.

Results: Since 1977, over 90 consensus panels have been held. A list of currently available OMAR reports is attached. The process and its results have been extensively researched (see reference list) and, in an earlier form, evaluated thoroughly by top-flight researchers such as Wortman and Kanouse. However, there is about a five to ten year gap between the process-as-it-is and the process-as-it-was-evaluated. Thus, the RAND study published in 1989 (five years ago as of 1994) looked at consensus conferences held under procedures of almost a decade earlier. The first of the nine separate evaluation substudies published before the 1989 final report appeared in 1986 ("Popular Press Coverage of Eight Consensus Development Topics" and "Treatment of Eight NIH

Consensus Development Conferences in the Biomedical Literature.") More recent third party evaluations of this magnitude are unavailable the RAND study cost about $1,000,000. OMAR conducts its own evaluations, however, for the CDCs and the program, and uses the results for program improvement. OMAR did not present these as part of available evaluation data, so I can not comment on their findings or technical characteristics. Further, there is a 1994 Department of Health and Human Services Inspector General report on how OMAR could improve.

If one believes that NIH followed many of the RAND recommendations and additionally has continued formative evaluations as a source of information, and if one believes this information has led to improvements, as I personally do, then the following evaluation data **systematically underestimate the value of the conferences as they are at present.**

- Consensus statement style varied, and included some findings so ambiguous, it was difficult for experts to figure out what "good" practice would be. (OMAR has since addressed this in its guidelines to achieve statements that are concise, didactic, practice-oriented, and clear.)

- About 30% of a nationwide sample of about 1,000 physicians knew about the Program; 50% said they had heard about the recommendations from at least one conference. About 14% to 30% of physicians who should have known about specific conferences in their area were sure they had, with high percentages among specialists and the highest, among physicians in Continuing Medical Education.

- **Timing was key. Attention peaked in a three month period and waned fast in the media and scientific press.** (OMAR has tried to tighten the connection between release of statements and speedier publication in key journals. All statements are now sent quickly to the Journal of the American Medical Association or to different but appropriate journals. Actual publication may take 5 or 6 months but this is beyond OMAR's control!)

- Citation analysis of the biomedical literature showed the statements had infused both the popular press and the scientific literature for six of the eight conferences in the sample selected for intensive evaluation.

- In the early period, some proceedings were published. RAND found that publication of the proceedings as well as the consensus statement itself promoted physician awareness. (OMAR staff indicate that the proceedings are not usually published.)

- o Physicians were more persuaded by information highlighting clinical, rather than research implications--practice-relevant information. They did not regard the statements as intrusive. (No action necessary)

- o By linking the physician survey with a hospital practices survey, RAND could test the link between awareness and practice. Both showed that the panels meaningfully influenced practice but more could be done. The levels of influence achieved might make the Department of Education jump for joy but RAND believed more could be achieved.

  In RAND's view, both substudies indicated the glass was about half-full (about 50% compliance with the consensus panel statement, with the statement on cesarean delivery having the greatest impact). Among the factors accounting for greatest impact were: the issue was widely perceived as a real problem; the conference was unique in drawing together accumulated information and stating its implications for practice--- there were no prior major reviews aimed at clinicians; and the statement was speedily printed in major obstetrical journals. RAND observes,

  > "Given a scientifically grounded and clinically relevant message delivered for the first time to a receptive audience that recognized the need for change, it is not surprising that this conference was more successful than most in changing physicians' and hospitals' practice." (Kanouse et al., 1989, p. xvii)

Thus, in terms of results, the consensus statements between 1978 and 1986 were informing physician practices to an extent RAND regarded as encouraging but with room for improvement. There are no more recent summative evaluation data through a study of the magnitude of the 1978 effort, but as noted, given the changes in OMAR practices, the current panels should be even more successful, both in raising the top and lifting the bottom for individual conferences.

OMAR can report, of course, on the number of consensus statements distributed, and other indicators of messages beamed. Further, OMAR officials are not necessarily wildly enthusiastic about the technical quality of the third-party evaluations. From my perspective, however, the studies seem remarkable in their thoroughness and efforts to put both qualitative and quantitive perspectives on panel process and results.

In addition, the Department of Health and Human Services, Office of the Inspector General (OIG) conducted a "replication" of one component of the earlier RAND study, picking up on the finding that continuing medical education programs were about the

most effective outreach for Panel findings. The 1994 OIG study was a mail survey of the population of chairs or directors of continuing education and the chairs of departments of family medicine, neurology, and oncology at all U.S. medical schools, supplemented with telephone interviews, discussions with NIH staff, examination of NIH materials, and discussions with experts (paraphrased OIG from 1994 report, p. 3). The reported response rate for Directors of CME was about 66%. The survey instrument, reproduced in the report, emphasizes "brand name familiarity" with the Consensus Program and specific recent Panel statements. The findings indicated greater familiarity with the Program and the Statements than earlier, particularly in the area of Oncology where 77 % of Directors reported holding a CME on the NIH Panel Recommendations, but---in that getting to be trite phrase, room for improvement. A few of those surveyed expressed concern with the "quality" of the OMAR process, such as time for reflection and the adequacy of panel selection. While the OIG report is careful to point out these may reflect a need for better information about the process rather than a change in the process, these concerns echo some of the major findings of the earlier Vinokur et al. (1985) and Wortman et al. (1988) reports, namely, "...selection bias, particularly with respect to the choice of questions and panelists." (1988, p. 469). This concerned Wortman et al. to such a degree that they raised serious question about the appropriateness of the NIH Panel Process for dealing with truly controversial, complex but medically significant issues.

Details of the RAND report, only briefly highlighted here, are worth thorough examination if OERI moves forward on consensus development panels, as are the Vinokur and Wortman papers and the 1994 OIG report. Dr. Kanouse and his colleagues would be excellent consultants although OERI should expect they will recommend (1) thorough study of how teachers, principals, and other potential audiences get the information on which they may act; (2) including in each proposal for a new topic strong evidence that there is a problem in actual practice (in one instances, the CDC belabored a practice largely abandoned before the panel met--the Halstead radical mastectomy); and (3) prospective evaluations monitoring the effects of individual conferences using fast, low-cost strategies.

And for more up-to-date information, if OERI moves forward on consensus development panels, the leadership and staff of OMAR would be available for consultation and discussion, just a few subway stops from OERI. OMAR has many visitors, both from the U.S. and from other countries, who are interested in the CDP and the panel processes, and stands ready to provide information and guidance.

THE AGENCY FOR HEALTH CARE POLICY AND RESEARCH CLINICAL PRACTICE GUIDELINES

Background: In an interesting example of Congressional concern with establishing research-based guidelines for medical practice, Congress established, in 1989, the Agency for Health Care Policy and Research (AHCPR). One of eight Public Health Service agencies, AHCPR evolved from the National Center for Health Services Research and Assessment. Among its' missions are facilitating the development of clinical guidelines, performance measures and standards of quality, and disseminating research findings and clinical guidelines.

There are two legislatively authorized mechanisms for guideline development: (1) convening panels of qualified experts and health care consumers and (2) contracts competitively awarded to public and nonprofit private organizations. Contractors have to appoint only AHCPR-approved panels.

Between 1992 and spring, 1994, AHCPR has released over eight clinical guidelines. Topics include acute pain management, urinary incontinence in adults, pressure ulcers in adults, cataract in adults, depression in primary care, sickle cell disease and management of cancer-related pain.

The 10 guidelines scheduled for 1994 release include low back pain, otitis media, heart failure, benign prostatic hyperplasia, treatment of adult pressure ulcers, post-stroke rehabilitation, diagnosis and management of unstable angina, and quality determinants of mammography.

Process: One of the driving forces behind the clinical practice guidelines is health care cost containment: the need to distinguish what is appropriate and necessary treatment of various conditions from what is inappropriate or unnecessary. Federal health care networks, such as Medicare, will not reimburse care found to be inappropriate or unnecessary.

Another driving force is the need to make decisions using the best possible scientific evidence even if a gold standard is not wholly met, through supplementing research with clinical experience.

Topic Selection Criteria: Both of these concerns are reflected in AHCPR criteria for selecting topics. According to Clinton et al. (1994), selection factors include:

- adequacy of scientifically based evidence on which to develop guidelines

- number of individuals affected by the clinical condition
- amenability to prevention
- expected potential for reducing inappropriate and unexplained variations in the presentation, diagnosis, management, or outcome of a particular disease or condition
- specific needs of the Medicare and Medicaid populations
- costs of the condition to all payers including patients

Panel Co-chair Selection Criteria: With regard to selection of the panel co-chairs (AHCPR has two leaders or co-chairs), once a topic is selected, organizations of health care practitioners and interested persons are invited to submit names. Criteria include, according to Clinton et al.,

- relevant training, clinical experience and leadership in their field
- demonstrated interest in quality assurance and research on the clinical condition, including publication of relevant peer reviewed articles
- commitment to the need for clinical guidelines
- recognition in their field with a record of leadership in relevant activities
- broad public health view of the utility of a particular procedure or clinical service
- demonstrated capacity to lead an interdisciplinary health care team in a group decision-making process
- demonstrated capacity to respond to consumer concerns
- previous experience developing guidelines for the clinical conditions in question
- no financial conflict of interest that would jeopardize the integrity of the guidelines

Panel Member Selection: AHCPR asks for nominations from a broad range of health care providers, including physicians, nurses, and consumers. Prior experience with developing guidelines for the condition is a big plus. Also considered is composing a

panel that has breadth and balance with regard to professional discipline, gender, minorities, and geographic regions. About 15 panelists are selected using criteria similar to those for the co-chairs.

Guideline Characteristics Sought: The Institute of Medicine helped develop guidance on attributes sought for the guidelines (Clinical Practice Guidelines, National Academy Press, 1990). According to Clinton et al., these include:

- validity: if guidelines are followed, they will lead to better health outcomes
- reliability/reproductibility: another set of experts given the same evidence would produce essentially the same statements and, in a clinical setting, varied practitioners interpret the guidelines the same way
- clinical applicability: guidelines are inclusive of the appropriately defined patient populations and should state to which populations they apply
- clinical flexibility: guidelines must identify the exceptions to the recommendations
- clarity: language should be unambiguous, terms clearly defined, and the presentation, logical and easy to follow
- multidisciplinary process: process for guidelines development must include participation by affected provider groups which includes but is not limited to serving on panels
- scheduled review: guidelines must include statements about when they should be revisited to see if the conclusions still hold. (AHCPR continues the panel with change in about 1/3 of the members to monitor new clinical information and determine when new guidelines are needed. AHCPR anticipates that guidelines will be revised every 18-36 months.)
- documentation: procedures used in developing the guidelines must be meticulously documented and described

Step-by-Step Method: According to Clinton et al., (1994) the guidelines are developed in an 11-step process:

1. Select clinical condition according to criteria
2. Define the clinical condition precisely

3. Review the scientific literature and available scientific evidence of appropriateness and effectiveness

4. Review estimates of outcomes important to patients who will be influenced by the intervention

5. Review benefits and harms from use of the intervention

6. Review health outcomes generated by the intervention

7. Review current and potential health care costs associated with the guidelines and **costs of alternative strategies for the prevention, diagnosis, treatment and management of the condition**

8. Invite comments on the guideline topic from professional and consumer organizations, researchers and manufacturers to be presented in an open meeting of the panel or in writing

9. Draft guidelines

10. Conduct external peer review of draft guidelines

11. Revise draft guidelines on analysis of pretesting and comments from external peer review

The AHCPR process may be of particular interest to OERI in the effort to utilize both research and clinical experience. Clinton et al. write,

> "The intention is to rely on scientific and empirical evidence as much as possible, but professional judgment and group consensus are also used in many steps of this process. Use of consensus will be necessary when insufficient empirical evidence exists to evaluate the effectiveness of treatments or procedures that are used in patient care. Nonetheless, these judgments must be carried out in an explicit manner." (1994, p. 33)

Dissemination: Each guideline appears in four formats, designed for different audiences:

- Guideline Report is the technical version including the guidelines and all supporting materials, primarily for use by researchers, educators and professional groups

- Clinical Practice Guideline is a short version intended as a desk reference for health care practitioners

- Quick Reference Guide for Clinicians highlights the Clinical Practice Guidelines for even faster reference by clinicians

- Patients' Guide is a consumer brochure on the condition and treatment options aimed at empowering the patient to become an informed partner in health care decision- making. **The consumer version is available in Spanish and English.**

State of the art information technology is used for dissemination, including print, on-line and electronic formats.

Evaluation: AHCPR's evaluation activities include, according to Clinton et al.

- managing the peer and pilot reviews of draft guidelines

- analyzing the costs of the recommended interventions

- monitoring receipt and awareness of guidelines

- comparing effectiveness of different modes of dissemination

- conducted targeted evaluation of specific guidelines in practice settings

- assessing the long-term effects of guidelines

- supporting investigator-initiated research on guideline evaluation

(AHCPR has an active research grants program; the Guideline development effort is embedded, institutionally, in this larger set of responsibilities for policy-related research.)

Time and Costs: Information on time and costs for each panel was not readily available in published descriptions I used. I estimate that the costs and time would be comparable to the OMAR process: that is, about $150,000 for each statement and about 12-24 months. The AHCPR process seems to involve more reviews of the guidelines after the panel meets, and more AHCPR involvement in the revisions.

Results: A formal third-party evaluation of the clinical guidelines program as managed by the AHCPR has not been published.

panel meets, and more AHCPR involvement in the revisions.

Results: A formal third-party evaluation of the clinical guidelines program as managed by the AHCPR has not been published.

Lessons Learned: In my opinion, the approach should be regarded as in development and somewhat experimental for the following reasons:

On Reason for Caution: AHCPR is drawing on and continuing to support a variety of approaches to developing clinical practice guidelines, approaches that vary almost all aspects of the process from criteria through how the consensus panels interact. One of the leading groups in this area, the RAND Corporation, has worked for more than a decade on development of guidelines for clinical practice.

> One researcher, Shekelle, is currently studying the important issue of guideline reliability. In the proposal for this award, Shekelle (1992) asks, "How would the results differ if different multi-specialty groups of experts were used to asses appropriateness and necessity?" (p. 30). Prior research has looked at guidelines developed several (or many) years apart. **Shekelle and his colleagues will be the first to systematically find out if similar guidelines are developed from two panels, using the same processes and with the same reviews of the literature, working at the same time.**
>
> In other words, the recommendations of the NIH panel for desirable qualities of Guidelines was more hopeful, than based on evidence that reliability, reproductibility, and validity can in fact be achieved. Some quite basic questions, such as the reliability of the guidelines, has not been systematically tested. Shekelle, citing a rich body of research in United States, the United Kingdom, Israel and the Netherlands, as well as many consortia, notes,
>
> > "In spite of all this activity, the reliability and validity of this method has not yet been sufficiently demonstrated. Since lack of an existing gold standard of appropriateness was the impetus for the development of consensus methods, no concurrent validity exists. No method has been systematically examined for prospective validity..." (p. 35)

Process: Because of the significance of their work, it seems useful to describe the RAND process in some detail. RAND's Health Services Utilization Study panel (now receiving AHCPR funding) "combines the feedback elements of the Delphi technique, the discussion format of focus groups, and use of expertise of Glaser's state-of-the-art

technique...[it] involves the rating of a detailed series of indications for appropriateness by a nine member panel of experts usually a formal group judgment consensus process. Two rounds of ratings are obtained, the first round via the mail and the second round after a face-to-face meeting." (Skekelle, p. 34)

Following topic selection, the next step is a search of computerized data bases, bibliographies and experts for all relevant literature on the procedure, particularly randomized controlled trials testing efficacy. The synthesis is sent for peer review by specialty societies.

From this literature and discussions with experts, RAND staff develop a list of indicators describing all possible clinical situations that might be encountered in as much detail as practical. The aim is to categorized patients into categories that are homogeneous for appropriateness.

Panel nominations come from appropriate specialty societies who may provide up to 10 nominations for each "slot". The requests usually yield about 80 to 100 candidates.

A nine member panel is selected to include a mix of academic and private practice physicians, from diverse parts of the country, and from specialties are patients likely to see. Physicians who perform the procedure are a minority.

The literature review, implications and instructions are mailed to panelists. They are asked to rate each indicator on a nine point scale of appropriateness, defined as having enough expected benefits for that type of patient to exceed expected risks by a wide enough margin to be worth doing.

Only panelists and the moderator participate although the room may include observers. The first round ratings have been returned and summarized and the summary provides the framework for the panel discussions. (Only the moderator knows individual ratings, and can use this information to draw out someone who may have been an outlier but is quiet during the discussion.) After discussion, each panelist re-rates each indication. The second round ratings are the final ratings, and the median rating determines the appropriateness category.

Results of the RAND approach: The method has been used at RAND to assess appropriateness and necessity of coronary artery angiography, upper gastrointestinal endoscopy, colonoscopy, coronary artery bypass graft surgery, cholecystectomy, carotid endarterectomy, abdominal aortic aneurysm surgery, percutaneous transluminal coronary angioplasty, cataract surgery, spinal manipulation, hysterectomy, and, recently, the care of an illness (asthma) rather than a procedure.

Pilot tests of the method are underway across the United States and in the United Kingdom, Israel and the Netherlands. Clearly, the research is shaping an approach that can build consensus on matters of practice where there may be more clinical expertise than scientific evidence. **Thus, this method as it evolves may be of more interest to OERI/ED than the consensus development conferences per se.**

Much remains to be done, however, in development, starting with systematic evidence of reliability: for example, would two different nine person panels drawn from the same pool of experts, held at the same point in time, with access to the same scientific literature, make essentially the same statements on medical appropriateness and necessity? How similar are the statements when the more costly face-to-face methods are used compared to processes that are less costly? According to Shekelle, most prior research on the reliability of expert panels deal with size, mix of practitioners, moderator effects, and the degree to which individual panel members are representative of their specialties.

Will consensus panel methods relying heavily on expert judgment carry the same weight as the OMAR processes? While guidelines have been published for the procedures listed above, a formal evaluation of effects of these panels on practice is not yet available. However, since both the Medicare system and private insurers do review claims for medical appropriateness and necessity of procedures, relevant use may be first by these health care review organizations and infusion into practice would be expected to be fairly rapid since failure to meet the guidelines would be challenged.

A Second Reason for Caution: It is clear that the risks of failure are non-negligible. The Environmental Protection Agency also has experienced some problems and discusses the risks in its guidelines on when to consider (and not consider) consensus panels for negotiated rule-making.

Developing guidelines when the research base may be limited (among other factors) is a developing art form. An example may help illustrate these.

In January 1994, the American Psychologist published a series of articles that demonstrate what can happen to consensus panel statements when the issues involve high stakes competition among professional groups, less-than-compelling research, and practice rather than changes in medical technologies: the case of the AHCPR depression guidelines (Schulberg and Rush, 1994; Munoz et al., 1994).

Schulberg and Rush present their interpretation of the decision-processes involved in this specific instance. They note, for example, that 39 separate literature reviews were commissioned by psychologists, psychiatrists, social workers, epidemiologist, internists

and other experts on the epidemiology, clinical manifestations, diagnosis and treatment of depression. Over 100,000 candidate abstracts were identified by MEDLINE which were "..carefully read" by reviewers and winnowed to about 3,500 pertinent studies. The article then discusses other decisions such as the definition of depression adopted made along the way, and summarizes the Depression Guideline Panel's recommendations. Perhaps the key to understanding the furor the guideline generated among psychologists comes in two statements in the Schulberg and Rush article:

> "...the psychologist's role will be that of case consultant about complex differential diagnoses rather than [routine] clinical assessor.."
>
> "The psychologist's role in treatment is...providing time-limited, depression-specific psychotherapy within the parameters of managed health care." (1994, pp.39-40)

In the second article, Munoz et al. hit the fan in outrage at the emphasis on pharmacology as the first line of treatment for depression by primary care physicians and by the way, the fact that psychologists have been too largely cut out of the action. Two aspects of the argument are of particular interest to the OERI use of consensus panels. The first is that the major quarrel is not with the technical statement which is seen as more nuanced and uncertain but how it has been translated into a clearer pharmacological emphasis in the clinician's guidelines. The second is with the research paradigms and limits to what can be established from the scientific literature, such as generalizability from specialized settings to primary care settings, the adequacy of the double-blind clinical trials of the drugs, looking at broader outcomes such as improved social adjustment in studies showing the efficacy of psychotherapy, and lack of adequate attention to interactions of treatment with ethnicity, gender, and developmental stages.

The bottom line: AHCPR is in the process of revising the guidelines less than a year after their release with the fullest involvement of psychologists and the American Psychological Association. (AHCPR routinely expects to revise the guidelines and asks the panels to determine whether the state of research or practice will require re-examination is 18 to 36 months or longer. The turn-around on the depression guides does not seem driven by breakthroughs in research or practice.)

Thus, if OERI is interested in the AHCPR process, more detailed information in this rapidly developing form can be obtained from Agency leadership and staff. However, in my opinion, the melding of clinical judgment and research evidence through the consensus panel process remains very much in evolution even in the health area.

NEGOTIATED RULEMAKING AT THE ENVIRONMENTAL PROTECTION AGENCY

Background: The Environmental Protection Agency's (EPA) negotiated rulemaking is a remarkable example of the consensus process in federal decision-making related to disseminatior. and policy. EPA has between 200 and 250 separate rulemaking activities underway at any one time. The conventional procedure is gathering data from various sources, informal meetings with affected group to try out ideas and get more information, drafting the proposed rule, publishing the draft for public comment, possibly holding hearings, and then preparing and publishing the final rule---a process that can sometimes take years.

Initiated in 1983 as a pilot project to expedite rule-making, EPA **in carefully selected circumstances**, now uses a consensus procedure called negotiated rulemaking (EPA, 1992).

Among the topics selected have been nonconformance penalties, pesticide emergency exemptions, farmworker protections, asbestos in schools, woodstove performance standards, small nonroad engine emissions controls, wood furniture manufacturing industry VOC emission controls, architectural and industrial maintenance coatings, disinfection by-products, national emission standards for coke oven batteries, oxygenated and reformulated fuels, recycling lead acid batteries, and fugitive emissions from equipment leaks. (EPA, 1993)

Process: Paraphrasing from EPA (1987), in negotiated rulemaking, the **appropriate EPA officials and representatives of groups affected by or interested in the topic** hold **a series of meetings** in which they try to reach consensus on all or most of the important features of the proposed rule. The negotiating committee **is chartered as an Advisory Committee under the Federal Advisory Committee Act.** Its plenary and subcommittee meetings are open to the public.

The process is governed by contractual understanding: on the part of the Committee to reach consensus if at all possible and on the part of EPA to abide by the results. **"To the extent that any consensus reached by the negotiating committee is consistent with the Agency's statutory authority, EPA is committed to using it as a basis of a Notice of Proposed Rulemaking."**

Criteria for selecting rulemaking topics: EPA wants to avoid making a tough situation worse in a failed negotiation. The EPA approach has evolved over time and the most

current guidance (1992) is reproduced in full as part of this appendix. For the purpose of linking evaluation results and the NRP to which they refer, however, criteria used earlier are discussed in some detail next (also paraphrased from EPA, 1987):

- There are a number of interrelated issues to be resolved, there are several ways in which they can be resolved, and relevant statutes can accommodate options. In other words there is more than one way to skin that particular cat that would be acceptable to EPA---as it were.

- The rulemaking will not challenge participants' fundamental values, so there are no deep-down emotional obstacles to reaching consensus.

- There is a statutory or judicial deadline for completing rulemaking **or some other action-forcing mechanism.**

- The prospective participants share some common ground on at least some of the issues to be negotiated.

- **The costs and benefits of the rulemaking are concentrated on a few entities.**

- There are relatively few interested and affected parties and they are readily identifiable.

- The prospective participants are willing to negotiate in good faith.

- The prospective participants see themselves as having an ongoing relationship with EPA.

That is: incentives exist to reach closure among a relatively small---but appropriate---group of stakeholders who are going to continue to work together, and with EPA. And vice versa in EPA's incentives.

After the proposed rule is published, rulemaking then proceeds in usual manner: the agency publishes the draft rule, considers public comments, prepares and publishes the final rule. The negotiating committees may be consulted informally but on-the-record about some public comments but play no formal role in the final rulemaking.

Time and Costs: The materials I read did not have detailed information but an evaluation by EPA's Office of Policy, Planning and Evaluation on the first seven uses (1987) indicated the process could be faster than the usual procedures. In this instance, absolute numbers are less meaningful than a comparative analysis of the negotiated versus the

usual rulemaking procedure. The fact that EPA has continued the approach indicates its value in circumstances meeting the selection criteria.

Results: The internal evaluation looked at how well the procedure worked and how it compared with conventional rulemaking. The well-designed study used documentary analysis (published literature, commissioned papers, minutes and summaries of the meetings, texts of the negotiated proposals, and other public documents) and interviews with EPA managers and staff members and with participants in all of the seven negotiated rulemakings. EPA noted that a good controlled comparison could not be made because the best substantive matches were developed at different times and under earlier regulatory situations. EPA apparently has not conducted a dual-track experiment at the same time with the same rule.

The evaluators concluded that:

- The negotiated rulemaking proposals (NRP) in numerous instances have been more pragmatic and could produce better environmental results than conventional rulemaking (CRM). As examples,
    - o More comprehensive and explicit definitions of what is meant by "no migration" for underground injection of hazardous wastes
    - o Larger area to be checked for the presence of hazardous wells than EPA had planned to propose
    - o A more accurate method of measuring woodstove emissions than EPA had originally suggested
- NRM has sometimes facilitated information exchanges and understanding of issues in dispute, both where consensus was reached and in some instances where it was not.
- Developing proposed rules through negotiation made final rulemaking easier and less costly for about half of the instances, with savings of $150,000 (1987 dollars) reported for the Woodstoves example. In this example, timeliness was as important as dollar savings. The evaluators note, "..it appears that it is easier to obtain EPA management concurrence in a negotiated rule than in one developed through the conventional process." (1987, p. 7)
- Working relationships developed during negotiations can transfer to constructive work in other situations

- However, the evaluators also noted that in these early examples, negotiations broke down in one instance (farmworker protection), only informal agreement on some of the issues could be reached in another instance (underground injection of hazardous wastes) and in a third instance (asbestos in schools) where limited agreement was achieved on many issues, public comments were critical on key issues.

Lessons Learned: Although a more recent evaluation has not been reported, the 1987 study offered some constructive suggestions which, if followed, should have increased the success rates. These include awareness that (1) NRM is labor-intensive particularly for the EPA managers who are the negotiators; (2) coordination within EPA and with OMB is more critical than usual since the EPA negotiator is speaking for the federal government; (3) there is no reduction in EPA time for preproposal data-collection and analysis; (4) NRM heightens expectations and failures may worsen EPA's relationships; (5) nonprofits are at a greater-than-usual disadvantage in terms of expertise. Maximizing success includes picking rules affecting program implementation rather than program structure; avoiding issues of controversial national policy or with complex multi-media implications; selecting issues where participants are likely to agree on common goals; selecting issues that can be resolved without setting wide precedents; selecting issues that do not intersect with ongoing litigation in a way affecting participants' ability to negotiate openly; and allowing enough time (four or more weeks) between negotiating sessions for reflection and consultations. And, say the evaluators, consider the risk of not achieving the objectives, likely consequences of failure to reach consensus, and whether alternative procedures would be better.

While this may not sound like a ringing endorsement for the NRP, EPA experience after 1987 has refined the procedures and has led to a clearly laid-out guide to selecting the best approach for consulting with external interests, including but not limited to negotiated rulemaking (EPA, 1992). EPA has identified a spectrum of consultation and consensus-building approaches, ranging from one-time information-exchanges through public hearings to the series of meetings involved in regulatory negotiation. They have identified clearly where the "FACA Line" is drawn, and developed detailed assessment criteria for selecting the appropriate consultation process that seem so useful for ED/OERI that the entire paper is reproduced as part of this description.

APPENDIX VI: BUT IS THIS FOR OERI? ANALYSIS AND RECOMMENDATIONS

Assume that experience both within the Education Department (in NIDRR for example) and elsewhere cumulatively suggests that consensus panels are at the least a promising practice for dissemination. The classic question in dissemination then is whether the processes apparently necessary for good results can be replicable, transferred, transposed, adapted, or modified---whatever--in a way that will be feasible for the adopting agency and achieve good bang for the scarce buck.

Two approaches were taken to answering this question: expert opinion (N= 1 expert in a formal interview and 20 other knowledgeable people in more informal discussions) and analysis of the match between what seems to be required for consensus panel effectiveness and the ED/OERI context. Both approaches yield the same result---an amber light, suggesting one proceed with caution, not a clear red or green.

POSSIBLY NOT--A PERSPECTIVE FROM AN EXPERIENCED EDUCATIONAL RESEARCHER: Many researchers within OERI, in other federal agencies and in the private sector (see Appendix I for a list of contacts) have had experience managing consensus building activities and some, managing consensus panels.

Dr. Richard Anderson, a national authority on reading and leader of the OERI-supported effort that led to "Becoming A Nation of Readers" has had recent experience with two consensus building activities. His observations (summarized below from notes on a discussion) present ideas others may express less trenchantly.

> How it is done is pretty important, Dr. Anderson observed. A consensus panel can promote an idea. However the process isn't always magic and can involve much bickering and dissemination and blanding out the report. It can be difficult if not impossible to get consensus in contrast to a process that obtains wide input but is not constrained by a demand for unanimity. That is, consultation is a much easier process than consensus, and may get also help to promote an idea.

Dr. Anderson gave an example of a consultation panel that worked, and a consensus panel that seems to be in difficulty.

> For example, "Becoming a Nation of Readers" was about 1 l/2 years in preparation. Dr. Anderson carefully constructed an advisory board that was representative of extremely diverse researchers...leading people whose conclusions would get a lot of

respect, lending credibility to the report, and who were reasonable. He also sought out all kinds of people throughout the entire process and in reviewing the drafts of the report. However, they were advisory, and did not have veto power. He listened, learned, considered, to achieve a consensus by keeping the ideas he believed were correct while changing ways of expressing those ideas that set teeth on edge...particularly on issues such as readability formulae and the meaning/phonics wars.

In contrast, the panel responsible for the National Standards in the Language Arts is battling once again on meaning/phonics plus lots of new battlegrounds waiting beyond beginning reading. These include whether to teach grammar; whether to teach spelling; and whether there should be a canon of literature, and if so, how to balance the White Anglo-Saxon Male authors with all other "perspectives" of authors. The issues are numerous and highly contentious; the panel has lots of field involvement; and the priority of building some kind of consensus on these emotionally charged issues has taken precedence over writing down a platform for the National Standards. Further, the authority is distributed: this panel does have the equivalent of veto power. Means for reaching any decisions are weak and limited. No systematic review of the research literature is involved. Overall, Dr. Anderson concludes this panel is in for a much rougher ride.

In discussing implications for OERI of these two approaches, Dr. Anderson observed that the issues in education may be less amenable to a consensus panel approach (in the National Standards sense, not a consensus building approach in the sense of "Becoming a Nation of Readers).

As Dr. Anderson sees it,

- On many educational issues, lines are drawn ideologically. Even leading researchers may not accept conclusions that disagree with their own beliefs, discrediting the research on which conflicting findings are based. In some instances, this may take the form of criticizing specific research studies. In other, much more difficult-to-deal-with instances, the entire research paradigm may be discredited, in Dr. Anderson's experience, on grounds such as the approach was not empowering to those studied.

  To a certain extent, valid criticisms are being raised. Behavioral and educational research lacks the measurement precision, reliability and validity of much research in the physical sciences, for example; few research studies are without their limitations which the researchers themselves often are the first to present; and the various research paradigms have strengths and weaknesses. However, the denial

goes far beyond recognizing these into refusing to accept results that are reasonably solid because they would mean giving up one's own position on an issue----such as bilingual education, phonics, a canon of literature, or tracking.

- o If one tried to find consensus but reach closure by permitting majority/minority opinions, the report might be published but have more value as an explication of the perplexities than as a guide to action. Announcing this option in advance might be an incentive not to find a negotiated position.

- o The third "out" is equally unattractive: blanding down. If the report in essence says nothing but platitudes, trivialities and the obvious, no one will disagree or act.

Among the other experienced educators, some were more optimistic at least on a pilot basis. Many expressed concerns similar to those of Dr. Anderson. Variations seemed in part related to what each meant by "consensus panel" and whether the expectation was primarily showcasing that educational research COULD lead to actionable recommendations to improve practice, or if the goals were closer to the visibility and general authority of the NIH-like processes with primary reliance on an adequate body of very high quality research.

POSSIBLY NOT--SOME ANALYTIC CONSIDERATIONS: Experts such as Ward (1980), Reid (1993) and the National Research Council Panel on OERI (1992) have recommended ED use of consensus panels in dissemination. The Office of Management and Budget examiner is said to have been mightily skeptical. Although reasons for the examiner's caution were not discussed, both general and specific experience with transferability of innovations suggest areas of similarity but some significant dissimilarities in context between OERI and other consensus panels.

Among the areas of similarity are:

- A common concern with bringing together research results to help improve practice and trying out new approaches to synthesis, to reaching conclusions, to making recommendations and decisions about effective innovations, and to influencing practice.

- A common investment, albeit of different portions, in knowledge creation by the agency

- A charter, mission, and expectation from Congress that the agency will fund research that sooner or later, and preferably within their lifetimes, will improve practice

- A history of various efforts to "get the message across" in dissemination including technical assistance, identification of effective practices (products, programs, policies), and diffusion of information through a wide variety of methods

- A history of research on criteria of effectiveness, including evaluation/research design, measures, and analysis, and elaboration of other aspects to be considered in identifying effective innovations

- Operating in a political context which includes many stakeholders with sometimes overlapping, sometimes conflicting agendas

- Dealing with expectations of relatively simple answers to questions that can be highly complex, where knowledge accretes slowly and ignorance does not always succumb to a frontal assault of demands or money thrown at it

- Knowledge that is almost always less than complete and less than perfect, where judgments do have to be made about conclusions based on strengths and limitations of a study

- A less-than-direct relation between a federal agency and practitioners, with much sensitivity about federal (or state) agents telling local practitioners what to do

- A broad professional culture of respect for the conclusions of leaders in the field, and of being influenced by major, convincing new knowledge

However, there are dissimilarities in the context. Among these dissimilarities--which argue against OERI use of consensus panels in the NIH/OMAR sense---are:

- The relationship of NIH to the health system is different in crucial ways from the relationship between ED and the educational system. NIH is supposed to find the shortest distance between life-preserving research and practice everywhere. ED is supposed to provide a knowledge bank on which local practice can draw if it chooses.

- There are no direct levers to practice. NIH panel decisions can trigger Medicaid payments, CHAMPUS practice, and Medicare system payments. Research affects the flow of money from ED mostly through the courts. For example, the civil rights decisions requiring children with handicapping conditions to be integrated as fully as possible in regular classrooms has led to a huge expansion of federal resources for local schools, in an effort to try to cushion some of the budget shocks of keeping up with court-mandated "rights and entitlements." One factor in the

court decisions was research suggesting special classes for handicapped children were perceived as stigmatizing, and were associated with less determined instruction and less development than the children's capabilities permitted.

- The health field has authoritative publications such as JAMA, which is widely read by physicians as setting standards for sound medical practice and knowledge. Apparently it is also widely read by lawyers, either directly or through the prime time and headline summaries. Education has nothing quite comparable.

- There is no educational malpractice business. Students don't take to the courts for damages due to ignorance that schools could have prevented the way patients take to the courts for damages in failures of medical practice. This may reflect the difficulty in proving the school was responsible (relative to the home or the child's own motivation) but there is not currently the stick of malpractice as well as the carrot of promoting learning involved in using every ounce of research to improve practice.

- In education, many points of practice can be so politically saturated (for example, bilingual education), rearranging old research doesn't stand much of a chance. It is major new research that can shake (if not shift) habits of thought, such as the Lazar/Darlington longitudinal studies on preschool and Coleman's analyses of private schools.

- ED, as an institution, may not have a culture of results; it may have a culture of process, new initiatives and response to new leadership. The supports required to take the highest scientific ground---such as staff who are themselves expert researchers, commitment to scientific integrity, ability to focus on scientific issues only without having to incorporate other important but not necessarily compatible goals---may exist fitfully more than consistently. (Demarest, 1994).

- The support and money required beyond technical assistance, research syntheses and high tech clearinghouses may not be available. Consensus panels are not cheap and the know-how to use them well can not be cranked up once every five years or so.

- ED, as an institution, may be strong on short-term memory rather than a four-year memory. By the time a formal consensus is in, using the OMAR example, the priorities and appointees may have changed several times. Topic setting may not be kept sufficiently neutral as new appointees look for discretionary money for their own agendas.

- What is high quality research in medicine is fairly well-established: double-blind, randomized clinical trials with sufficient numbers to assess what works for which patient. What is high quality research in education is still evolving, although within the qualitative and quantitative paradigms, there are standards of quality.

- However defined, there may not be a great deal of high quality research for OMAR type consensus panels. The situation might be closer to the appropriate practice panels (RAND) which rely more on clinical experience to supplement quality research.

- NIH/OMAR topics are solution oriented: for example, whether ulcers are primarily caused by phylobacteria more than stress and whether these succumb to bad bacteria zappers effectively and safely; whether cochlear implants improve hearing notably and safely for certain types of deafness. OERI/ED topics, judging by the titles of knowledge-into-practice papers, are problem oriented: for example, the reasons for problems in beginning reading and what research says broadly about how to teach new readers such as the importance of motivation and of comprehension rather than decoding alone.

APPENDIX VII: A TALE OF TWO LOOKS AT RESEARCH SYNTHESES

Several ED colleagues sent copies of research syntheses they saw as relevant to consensus panels. Some were developed from ED-sponsored conferences such as the 1990 national conference on student motivation intended to bring research and craft knowledge together to guide practice. Some were reports of ED-sponsored activities to guide both practice and research, such as the publications emerging from the National Forum on Educational Statistics. Others were research syntheses aimed at helping teachers and principals use research to improve instruction, such as the 1980 Research-Within-Reach series on mathematics and the 1992-93 series on Becoming a Nation of Readers.

Because the main audiences for these publications were practitioners, they are short on descriptions of process. In some instances, interviews have helped identify a consensus panel effort from a consensus-building effort. Where possible, these have been cited in the main text. However, because of the sketchiness of information on issues such as how the topics were chosen as amenable to a synthesis, how the panels and speakers were selected, and whether an effort was made to shake a topic until consensus was achieved or the fundamental disagreements revealed, some highly innovative and relevant ED experience obtained through preparation of these syntheses may not be included in the text.

What seems clear is that obtaining input from experts is standard operating procedure for ED. Generally, based on this array of instances, ED routinely uses input and consensus building techniques, often relies on the skill of such talented and experienced conference leaders as Tomlinson to capture the essence of meetings, and goes to great lengths to get extensive reviews of draft reports by federal, state and local educators and researchers.

These approaches seem to have served well in the past, and perhaps there is little room for improvement. A few observations may be helpful, however, with the "keep up the good work" items first.

1. Cheers on reader friendliness! Looking at the array spanning almost 15 years, the most recent publications seem like designer genes: made to fit the questions, interests, preferred style, credible authorities of segmented audiences. The effort made to keep the writing jargon-free shows; and the layouts use shaded blocks, italics, white space, real-people quotes and vignettes and other devices most niftily to keep those pages turning.

2. And cheers on publication flow! About 15 years ago, one had a sense of a mine of stuff that might help improve practice waiting to be exploited. Looking at these documents, one has a sense that whatever research was available was being squeezed for

all it was worth to help answer the enduring and more recent questions about educational practice. If all the ERIC Clearinghouse reports, the publications from the Laboratories and Centers, and from ED itself aimed at dissemination were put together, the annual list probably would be as large as most folks would be able to keep up with...and pretty well targeted on the issues of the day.

However, three questions arose:

1. Does ED have any "minimum daily requirements" for its research syntheses? Except for the glossy covers, admirable white-space readability, and common print style, the publications vary tremendously in areas where it counts most, such as how and how much knowledge is guiding recommendations and the extent to which the conference was preaching to the converted or trying to rustle agreement among groups with major disagreements on the issues.

Most publications appear to rely on commissioned papers, which were discussed and surely revised; but the guidance given the researchers is uncertain with regard to issues such as years to be covered, breadth of the literature to be examined (U.S., international? through what disciplines?) and methodology to be used (wise mind? meta-analytic? other synthesis techniques?). Further, some publications include some references to specific research while others may be based entirely on craft knowledge as far as documentation is concerned.

Having some methodological standards for research synthesis and making public what level(s) or approaches were used might increase reader understanding of the comprehensiveness and quality of what's between the covers.

2. Does ED have any expectations about looking beyond the educational canon when going from research to practice? Comparing a 1993 symposium on reading research published in Psychological Science with the ED reading series, I found no overlap in citations and virtually none in sources. The ED syntheses drew on sources such as the Journal of Educational Psychology, Reading Research Quarterly, and Journal of Reading Behavior; reports of commissions; publications of the Laboratories and Centers; prior ED reports; and books specifically on instruction. The Psychological Science issue drew primarily original sources such as the Journal of Experimental Psychology, Psychological Review, and Science; and more rarely, on books such as "The Biology and Evolution of Language".

It doesn't seem much of a feather in Psychological Science's cap that the research literature on reading emerging from educational studies should be so ignored, but it isn't much of a feather in ED's cap either that there appears to be no expectation that research

outside of the educational canon be examined. Some ED publications may do exactly this, and just didn't happen to be included in the array.

At a minimum, in ED all publications aspiring to bring together the best of current research on a topic (be it math, reading, learning how to learn, motivation---whatever) ought to examine thoroughly the relevant research outside of education per se.

3. Does ED have any process for cumulating knowledge across publications? ED publishes knowledge syntheses on the same topics within a few years of each other. However, no one seems to be looking back to cumulate knowledge. For example, the terminology, recommendations and information in publications as the 1993 "Transforming Ideas for Teaching and Learning to Read" didn't seem to build on the 1989 "Becoming a Nation of Readers."

Sure, someone could sweat out a cross-walk between closely related syntheses, such as the ED/Clearinghouse on Early Childhood synthesis on school readiness and ED's own 1994 synthesis of research on school readiness, but why make it difficult? If more recent studies confirm earlier conclusions, that's worth saying. If older approaches have been superseded, make clear what the newest evidence shows. Why give the impression of either nothing new has been learned or no one reads any report besides the one they are writing, an impression that surely is not true.

De minimus, editors and writers could systematically check all prior ED knowledge syntheses and publications on a topic to cross-walk recommendations, constructs, language, etc. If possible, the ED syntheses could cumulate knowledge, terms, findings, recommendations rather than appearing so helter-skelter.

References to Appendices VI and VII

Binkley, Marilyn R. Becoming a Nation of Readers: What Principals Can Do. Washington, D.C.: U.S. Department of Education, 1989.

Binkley, Marilyn R. Becoming a Nation of Readers: Implications for Teachers. Washington, D.C.: U.S. Department of Education, PIP 86-400, 1986.

Driscoll, Mark J. Research Within Reach: Elementary School Mathematics. Washington, D.C.: U.S. Department of Education. April 1980.

Fiez, Julie A. and Peterson, Steven E. "PET as Part of an Interdisciplinary Approach to Understanding Processes Involved in Reading." Psychological Science, 4 (5), September 1993, 287-293.

Forgione, Jr., Pascal D. and Martin E. Orland. A Guide to Improving the National Education Data System. Washington, D.C.: U.S. Department of Education, NCES 91-140, 1991.

Gernsbacher, Morton A. "Less Skilled Readers Have Less Efficient Suppression Mechanisms," Psychological Science, 4 (5), September 1993, 294-299.

McKinney, Kay. Improving Math and Science Teaching. Washington, D.C.: U.S. Department of Education. PIP 93-1005, 1993.

Means, Barbara, John Blando and Kerry Olson. Using Technology to Support Education Reform. Washington, D.C.: U.S. Department of Education, OR 93-3231, 1993.

Posner, Michael I. and Bruce McCandless, "Converging Methods for Investigating Lexical Access," Psychological Science, 4 (5), September 1993, 305-310.

Raney, Gary E. and Keith Rayner. "Event-Related Brain Potentials, Eye Movements, and Reading." Psychological Science, 4 (5), September 1993, 283-287.

Rayner, Keith. "Reading Symposium: Introduction." Psychological Science, 4(5) September 1993, 280-283.

Sashkin, Marshall and Egermeier, John. School Changes Models and Processes: A Review and Synthesis of Research and Practice. Washington, D.C.: U.S. Department of Education, PIP 93-1108, 1993.

Seidenberg, Mark S. "A Connectionist Modeling Approach to Word Recognition and Dyslexia," Psychological Science, 4 (5), September 1993, 299-305.

Stein, Marcy. The Beginning Reading Instruction Study. Washington, D.C.: U.S. Department of Education, OR 93-3153, 1993.

Sweet, Anne P. State of the Art: Transforming Ideas for Teaching and Learning to Read. Washington, D.C.: U.S. Department of Education, OR 93-3046.

Tomlinson, Tommy. Issues in Education: Hard Work and High Expectations: Motivating Students to Learn. Washington, D.C.: U.S. Department of Education, PIP 92-1500, 1992.

Office of Medical Applications of Research

OMAR Fact Sheet

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES | Public Health Service | National Institutes of Health

The Office of Medical Applications of Research (OMAR) is the focal point for health technology assessment and transfer activities at the National Institutes of Health (NIH).

Located in the Office of the Director, OMAR works closely with the various NIH Institutes, Centers, and Divisions in an effort to improve the process of translating the results of biomedical research into knowledge that can be used effectively in the delivery of health services. Additionally, the office is involved in evaluating the safety and effectiveness of drugs, devices, and medical procedures that are already in general practice.

A major responsibility of OMAR is the coordination of the Consensus Development Program, through which medical technologies are assessed by non-Federal specialists, generalists, other health professionals, and consumers in open forum and the results disseminated widely to the health care community.

Another area of responsibility for OMAR is the coordination of NIH responses to Medicare coverage issues raised by the Public Health Service (PHS). Specifically, the PHS asks NIH experts to evaluate safety and effectiveness of drugs, devices, and procedures that are being reviewed for possible reimbursement under Medicare.

In carrying out its overall mission of technology assessment and transfer, OMAR:

- Works with NIH Institutes, Centers, and Divisions to promote effective dissemination of information gathered through consensus conferences and other assessment activities.
- Advises the NIH Director on development in medical technologies and the applications of medical research.
- Provides a link among the technology assessment activities of the Institutes, Centers, and Divisions of NIH.
- Monitors the progress and effectiveness of NIH health technology assessment and transfer efforts.

Additional information about OMAR and its activities may be obtained from the Office of Medical Applications of Research, Building 1, Room 260, Bethesda, Maryland 20892, phone (301) 496-1143. Copies of past Consensus Statements as well as the schedule of upcoming Consensus Development Conferences also are available.

NIH Consensus Development

U.S. Department of Health and Human Services

Public Health Service

National Institutes of Health

Office of Medical Applications of Research

The National Institutes of Health launched a program in 1977 designed to improve the lines of communication from the health research community to the practicing physician and the public. The key element in this effort is "consensus development," a process that brings together biomedical research scientists, practicing physicians, consumers, and others in an effort to reach general agreement on whether a given medical technology is safe and effective. That technology may be a device, drug, or medical or surgical procedure.

The Consensus Development Program is aimed at complementing—but not replacing—the usual means of reporting research results through publication in scientific journals and other medical periodicals and through the lay media.

NIH initiated the program because there was no formal process within the research community to assure that medical research discoveries were identified and evaluated to determine if they were ready to be used by doctors and other health care workers. Since NIH is the nation's principal health research agency, it was felt that it should assume the responsibility of more fully reporting biomedical research findings to the practicing community and the public.

In recent years, there has been considerable public criticism voiced concerning the use of certain surgical and medical procedures, drugs, and devices. Many have claimed that some new technologies have reached the health care delivery system without being tested adequately. On the other hand, there are those who maintain that some well-validated technologies have been too slow in making their way from the research work bench to the hospital bedside.

A main objective of the NIH Consensus Development Program is to provide the physician and the public with current, responsible information on the pros and cons of medical technologies. This information is made public through reports containing conclusions and recommendations about a given technology, written by expert and lay members of consensus development conference panels.

With the highly complex work associated with biomedical research has come rather technical language that is not always easily understood by all audiences. Consensus development panels have, therefore, worked to produce reports that, although appropriate for the practicing physician, will also be comprehensible to the general public.

The value of these reports is that they may identify safe and useful emerging medical technologies and make a wider audience aware of their availability. At other times, they may point out some potential problems that could result from the use of an existing technology. In some instances, panels may even recommend against using a given medical or surgical procedure, device, or drug, under certain conditions.

One of the prime objectives of consensus development conferences is to provide a public forum to ensure that all points of view are aired. Specific time periods are set aside at every consensus development conference to enable individuals and groups to raise questions or issue comments, and meeting summary reports are designed to include the different viewpoints voiced at the meeting.

For further information about the NIH Consensus Development Program, contact the Office of Medical Applications of Research, National Institutes of Health, Federal Building, Room 618, Bethesda, Maryland 20892, 301-496-1143.

Guidelines for the Planning and Management of NIH Consensus Development Conferences

OMAR Consensus statements are available from the NIH Consensus Program Information Service, PO Box 2577, Kensington, MD 20891

Tel. 1-800-644-6627 (1-800-NIH-OMAR)

and through the Internet via Gopher (gopher.NIH. gov) or World Wide Web (http://text.nlm.nih.gov)

NATIONAL INSTITUTES OF HEALTH
Office of the Director

Introduction

The Consensus Development Conference (CDC) Program is an influential and innovative activity in health technology assessment and transfer conducted by the National Institutes of Health (NIH). Under this program, which has operated since 1977, NIH organizes major conferences that produce consensus statements on important and controversial topics in medicine. Each conference is jointly sponsored and administered by one or more Institutes, Centers, or Divisions (ICD's) of NIH and by the Office of Medical Applications of Research (OMAR) in the Office of the Director of NIH. Depending on the topic, other Federal agencies with biomedical components may join in the sponsorship of a CDC.

The CDC Program complements other NIH activities in technology assessment and transfer. These include scientific publications, lectureships, symposia, workshops, conferences, colloquia, task force and committee reports, staff studies, and brochures developed for a variety of audiences including research scientists, health policymakers and providers, health educators, and the public.

The purpose of a CDC is to evaluate the available scientific information on a biomedical technology and to develop a consensus statement that advances understanding of the technology or issue in question (assessment) and that will be useful to health professionals and the public at large (transfer). A broad-based panel listens to the scientific data presented by experts, weighs the information, and then composes a consensus statement that addresses a set of questions previously posed to the panel. This statement is an independent report of the panel and is not a policy statement of NIH or the Federal Government. The panel is not an advisory body to NIH, although NIH or any other Government or non-Government organization may adopt all or part of the panel's recommendations.

CDC's examine either emerging or established technologies. Although many aspects of a technology under review may be discussed—economic, sociologic, legal, and ethical—the primary purpose of a CDC is to provide scientific evaluation. To this end, the statements in general and the recommendations

issued in particular should focus on medical safety and efficacy, although they may refer to other issues in passing. CDC's are particularly useful for providing guidance when a controversy exists in differing therapeutic or diagnostic options and when the issue is of public as well as professional interest. The timing of the conference should neither be so early in the developmental course of a new technology that data are insufficient nor be so late that the conference merely reiterates a consensus already reached by the profession.

A consensus statement is based on publicly available data and information. It is not intended as a legal document or as a primary source of detailed technical information such as one might find in a review article or task force report. Rather, the statement reflects the unified view of a panel of thoughtful people who understand the issues before them and have carefully examined and discussed the scientific data available on these issues. The creative work of the panel is to synthesize this information, along with sometimes conflicting interpretations of the data, into clear and accurate answers to the questions posed to the panel. When consensus cannot be achieved, the statement should reflect this by noting uncertainties, options, or minority viewpoints. Following the conference, the consensus statement receives wide circulation through both the lay and the medical media. The conference proceedings may be published and telecast, and a summary videotape may be produced.

The basic principles governing the conduct of a CDC follow:

1. A broad-based, non-Government, nonadvocacy panel is assembled for each CDC to give balanced, objective, and knowledgeable attention to the topic. Panel members are carefully screened to exclude anyone with scientific or commercial conflicts of interest (see Selecting the Panel).

2. The panel meets in public session for presentation of all data by invited experts, commentary, and discussion, and in executive session for preparation of the consensus statement.

3. A number of specific questions determine the scope and direction of the conference. These questions are developed

in advance, widely circulated, and known to all participants at the conference. The principal job of the panel is to develop responses to them.

4. At the close of the conference, the draft consensus statement is prepared by the panel in executive session and is presented in plenary session. Following public discussion and any revisions deemed appropriate by the panel, the statement is adopted formally and stands as the record of the conference.

5. Wide dissemination of the consensus statement is sought in an effort to achieve maximum impact of the statement on health care practice.

The Topic

Topics for CDC's may be suggested by the ICD's, OMAR, other government health agencies, Congress, and the public. Final selection of a topic is made when agreement is reached between a sponsoring ICD and OMAR. Sponsorship by additional ICD's or non-NIH organizations will be considered subsequently.

A topic must meet the following selection criteria:

- It should have public health importance; it should affect or broadly apply to a significant number of people.
- Controversy or unresolved issues should surround biomedical/scientific aspects of the topic that would be clarified by the consensus approach, or there should be a gap between current knowledge and practice that a CDC might help to narrow.
- It must have an adequately defined and available base of scientific information from which to answer the conference questions and to resolve the controversies insofar as possible.
- It should be amenable to clarification on technical grounds, and the outcome should not depend mainly on the subjective judgments of panelists.

The following additional elements are desirable for positive consideration of a conference topic:

- Health care cost impact
- Preventive impact
- Public interest.

After a topic meets the selection criteria, the planning and implementation of the CDC may proceed. Because of the variable subject matter and sponsorship of CDC's, no single, fixed format is appropriate. As long as the principles enumerated in the Introduction are observed, the format of a conference may be designed to best accommodate the topic under consideration.

The Staff

OMAR and ICD staff are responsible for planning and conducting the conference logistics, publicity, evaluation, continuing medical education accreditation, and any other tasks required to make the conference a success.

The sponsoring ICD nominates the ICD coordinator, who usually chairs the planning committee and serves as the representative of the ICD in managing the conference. The ICD coordinator should be knowledgeable in the area of science under consideration.

OMAR assigns a senior staff member to serve as the OMAR coordinator. This individual works with the ICD coordinator and other ICD representatives in organizing the conference and also serves on the planning committee.

The sponsoring ICD's information office assigns a staff person to assist the OMAR Director of Communications in planning conference publicity and media coverage as well as dissemination of the consensus statement following the conference. Both of these individuals serve on the planning committee.

The Initial Planning

The ICD coordinator, the OMAR coordinator, and other appropriate ICD and OMAR staff review the conference topic, review the general scientific base supporting the topic, and agree on whether the topic meets the criteria for a CDC. After acceptance of the topic, they outline the general conference objectives and establish a timetable for conference planning. The coordinators identify other interested organizations within and outside NIH and consider appropriate roles for them, which may include cosponsorship of the conference. Representatives from other ICD's and organizations may be asked to join the planning process at this stage.

An informal organizational meeting is held between the OMAR coordinator, the ICD coordinator, and representatives from the other interested organizations. Here the scope, title, and approximate date of the conference are tentatively determined, and a panel chairperson is chosen (see The Panel Chairperson). The panel chairperson takes part in all major conference planning activities. Additionally, two or three persons from the general research community, who are not Federal employees and are knowledgeable about the conference issues, are selected to become members of the planning committee.

At the organizational meeting, the participants might want to consider development of background review papers, meta-analyses, or use of a decision model to assist the panelists in dealing with the data presented during the consensus conference. Planning for evaluating conference impact is encouraged. The potential impact of the conference should be defined in terms that are amenable to quantitative measurement and evaluation.

The coordinators develop the travel and conference budget. The OMAR coordinator arranges logistical support for the CDC through a non-Government logistical support contractor (see Logistical Support).

103

The Panel Chairperson

The panel chairperson should be a knowledgeable and prestigious figure in the field of medical science under consideration but should not be identified with strong advocacy of the conference topic or with research that might be presented to answer any of the conference questions.

The individual selected will be responsible for chairing the plenary session of the CDC, the deliberations of the panel, and the press conference, and thus should be both a strong public moderator and a skillful leader of small group discussions.

The panel chairperson must be a citizen of the United States and must not be a Federal employee.

The Planning Committee

The planning committee has four major functions: to draft the conference questions, to draft the conference program, to recommend conference speakers, and to recommend panel members.

The planning committee is ordinarily composed of the panel chairperson and representatives from the sponsoring ICD, OMAR, and other interested Federal agencies within and outside the NIH. The planning committee also includes several recognized experts from the research community who are not Federal employees.

The planning committee is usually chaired by the ICD coordinator, although the ICD director or another representative may serve in this role.

To prevent the appearance of bias, no planning committee members, except for the panel chairperson, may serve on the panel. Planning committee members may serve as speakers at the CDC. Disclosure of scientific bias and commercial conflicts of interest is requested of planning committee members for the record.

The Planning Committee Meeting

The planning committee convenes a meeting for the purpose of performing its four major functions.

Drafting the Conference Questions

The agenda of a CDC is structured around key questions posed to the panel. Ordinarily, four to six questions are posed, including questions on efficacy, risks, clinical applications, and a final one soliciting the panel's opinion on directions for future research. These questions determine the scope and substance of the conference. They should be framed so that answers can be derived from scientific information presented by the speakers. The phrasing of the questions should not allow for responses based solely on subjective judgments or opinions of the panel. Questions should be straightforward and concise so that it will be evident whether consensus has been achieved.

Drafting the Conference Program

The CDC usually begins with a preliminary panel orientation and discussion on the day before the conference. The CDC then proceeds for the next 2½ days. The first day and a half consist of a plenary session in which speakers present evidence followed by open discussion among panelists and audience. On the evening of the first day, the panel meets in executive session to begin to draft the consensus statement. At noon the second day, the panel again meets in executive session and finishes drafting the consensus statement. The following morning the statement is presented publicly, modified at the discretion of the panel on the basis of comments from the audience, and adopted formally by the panel.

The planning committee should plan the entire program, including evening and luncheon executive sessions and alternate activities for the second afternoon. An appropriate balance should be provided among the times allotted for speakers, discussion, breaks for scientific exchange, and drafting the consensus statement.

The conference program should be broad enough to encompass the body of scientific information essential to answering the conference questions and should include divergent scientific and medical views. Presentation topics should be selected so that speakers can use style and format appropriate to scientific meetings; that is, the presentations should be based on analysis of data with methodology fully explained and citations to the relevant scientific literature provided. The intent is that speakers present information to the panel and the audience as scientific experts, not as advocates for particular answers to specific questions.

In some instances, the usual 2½ day CDC format may be insufficient for the creation of a thoughtful consensus statement or else unnecessarily long. Alternative formats include

- Providing the panel with written papers and/or position statements well ahead of the CDC
- Shortening the time allotted to speakers
- Extending the CDC for an additional day
- Holding more than one panel preliminary meeting.

Selecting Conference Speakers

Speakers are selected for their scientific expertise and may include clinical investigators and basic scientists as well as general authorities in the field. Where differences of scientific opinion exist, care should be exercised to include the presentation of opposing data and interpretations. Speakers are asked to confine their presentations to the scientific topic that they have agreed to address and to be certain to present all relevant data and information. Speakers are expected to provide abstracts of their presentations to the ICD coordinator at least 2 months before the conference.

Selecting the Panel

A range of expertise on the panel is important to the panel's ability to deliberate on the varied scientific material presented. This diversity enhances the credibility of the consensus statement. The panel should represent various sectors of

professional and community life, including each of the following four general categories:

- Research investigators in the field, i.e., basic and clinical scientists who are active in the area of consideration
- Health professionals who are users of the technology, including practicing physicians, dentists, psychologists, nurses, or other health care providers
- Methodologists such as epidemiologists and biostatisticians
- Public representatives such as ethicists, lawyers, theologians, economists, public interest group or voluntary health association representatives, consumers, and patients.

Panel members should not be professionally identified with advocacy or promotional positions with respect to the conference topic or with research that might be presented to answer any of the conference questions. Also, panelists should not have financial interest in the issues to be resolved.

Panel members must be thoughtful, able to weigh evidence, and capable of collaborative work.

The size of the panels has varied from 9 to 16 members, with 12 or 13 found to be a reasonable working group.

Because the CDC is an exercise in peer review, in which the panel reviews information presented by the experts, panelists should not be speakers.

Full-time Federal employees may not serve on panels so as to avoid the appearance of undue Federal involvement. Panelists must be citizens of the United States.

Conference Evaluation

The OMAR and ICD evaluation offices plan evaluation protocols and collect data both before the conference and at intervals following the conference. Every effort is made to collect input from planning committee members regarding techniques for evaluating the impact of the conference.

Conference Publicity

Planning of conference publicity is coordinated with the OMAR Director of Communications and the information office representative from the sponsoring ICD.

Planning includes development of an information dissemination plan that:

- Identifies target audiences thought to have an interest in the CDC so that an announcement of the meeting may be disseminated as widely as possible to all interested groups in a variety of media (e.g, professional journals, lay newsletters)
- Encourages interested parties to attend the conference and present testimony
- Addresses the possibility of a live telecast of conference proceedings
- Outlines the publication of reports and papers and preparation of videotapes or audiotapes resulting from the conference
- Identifies the organizations and publications to which the consensus statement should be sent.

Continuing Medical Education Credit

The NIH Office of Education is accredited to sponsor continuing medical education (CME) for physicians in Category I of the Physician's Recognition Award of the American Medical Association. Most CDC's average between 13 and 15 CME credits.

Occasionally, the potential audience for a CDC includes groups other than physicians (e.g., psychologists, dentists, nurses). Thus other types of continuing education may be appropriate for a conference. Often, the application process for other types of continuing education is handled directly with the sponsoring professional society (e.g., Academy of General Dentistry, American Psychological Association).

The Consensus Development Conference

On the day before the conference, the Director of OMAR briefs the panel on the consensus development process and procedures. Panelists may then discuss the areas in which they wish to probe individual speakers and outline their approaches to the conference questions.

The conference plenary sessions are moderated by the panel chairperson. The chairperson ensures that speakers adhere to time limits, allows ample opportunity for scheduled discussion, and invites comments from panelists and the audience, taking questions from panelists first.

The entire panel, working in subgroups, begins to draft the consensus statement during the evening executive session on the first night. When the noon executive session begins on the second day, the chairperson directs the panel to finish its statement within the allotted time. (Experience has shown that several drafts of the statement are necessary.) The panel should start this session with at least a first draft to work on.

The panel should attempt to reach consensus on each question based on the scientific evidence presented. To produce a firm statement, the panel is encouraged to draw conclusions and form recommendations whenever feasible. If consensus cannot be achieved, minority or alternative views should be included. A straightforward expression of diversity is preferable to a contrived or ambiguous consensus.

The conference convenes again in plenary session on the morning of the final day, and the proposed consensus statement is read by the chairperson. At this time, the statement is subject to review and comment by the conference attendees. The panel adjourns following this session for a final, brief executive session (approximately 2 hours long), where it may revise the draft statement to reflect comments from the floor. The first 10 to 15 minutes of this final executive session are spent on reviewing the draft press release, which highlights the conclusions of the panel as set forth in the consensus statement. The press release is written by the OMAR Director of Communications and the ICD information office representative and is distributed at the press conference and disseminated nationwide after the press conference.

After the final executive session, a press conference is held, with the panel chairperson serving as spokesperson for the panel. Panel members are expected to remain for the press conference, which usually lasts about an hour. The panel chairperson opens the press conference with a brief summary (about 5 minutes long) of the panel's findings in lay language. Subsequently, the floor is opened to questions from the media. The panel members and chairperson should participate in responding to press questions. The panel chairperson, in particular, as well as the other panel members should be available for a brief time after the press conference for possible questions and interviews by newspaper, radio, or television reporters.

Following the conference, the consensus statement is promptly edited for style, syntax, and clarity by the OMAR and ICD coordinators and sent to the panelists for their final review. The edited statement is then sent to the chairperson for final approval. Following this process the statement is considered final.

As soon as possible following the conference, the OMAR Director convenes a debriefing (followup) meeting for OMAR and ICD staff to review the strengths, weaknesses, and outcome of the conference.

Conference Results

CDC's usually receive considerable attention from the media. The nation's major print and broadcast media often report the results of a CDC, which serves to focus attention on the topic and the statement of the panel.

The OMAR Director of Communications and the ICD information office representative develop an information dissemination plan covering the publicity for the conference, the media coverage of the conference results, and the strategy for distributing the consensus statement.

The consensus statement is printed by OMAR and distributed routinely to a variety of Federal health agencies, health care organizations, directors of continuing medical education at

American Hospital Association member hospitals, deans of medical schools, directors of State and county medical societies, and directors of HMO's and PPO's. Additionally, the consensus statement is sent to targeted individuals and organizations specified in the information dissemination plan.

The Journal of the American Medical Association publishes many of the consensus statements, as do specialty journals in the area of the topic. In addition, OMAR places notices in numerous professional journals announcing the availability of the consensus statement and inviting inquiry. On several occasions, the consensus statement along with selected papers from a CDC have been published either as a supplement to a specialty journal or as a monograph.

Summary videotapes and audiotapes of the conference may be prepared and distributed to specialty groups.

A summary of the statement is also prepared and sent to appropriate specialty journals, popular medical and health magazines and newsletters, and health and science sections of major newspapers.

Logistical Support

OMAR maintains a logistical support contract to provide travel and management assistance throughout the planning and conduct of a CDC. The contractor assigns an individual to work with the ICD and OMAR coordinators from the beginning of the planning process through completion of the conference to the final distribution of the consensus statement.

Expenses for planning committee members, panel members, and conference speakers are paid by OMAR through the contractor.

The CDC budget developed by the ICD and OMAR coordinators governs the expenditure of funds from the logistical support contract.

For additional information on the NIH Consensus Development Program, contact the Office of Medical Applications of Research, NIH, Federal Building, Room 618, Bethesda, Maryland 20892, phone 301-496-1143.

AHCPR fact sheet

Agency for Health Care Policy and Research — April 1993

2101 East Jefferson Street, Rockville, MD 20852

Agency for Health Care Policy and Research

The Agency for Health Care Policy and Research (AHCPR) is a component of the U.S. Public Health Service (PHS). Established in 1989, AHCPR's purpose is to improve the quality, appropriateness, and effectiveness of health care, and to improve access to health care services.

AHCPR carries out its responsibilities by supporting and conducting health services research—including medical effectiveness research—and assessments of health care technologies. AHCPR also facilitates the development of clinical practice guidelines.

AHCPR Programs

AHCPR's research programs examine the availability, quality, and costs of health care services and ways to improve the effectiveness and appropriateness of clinical practice, including the prevention of disease.

Availability, quality, and costs of health care services

AHCPR research enables policymakers to better understand and monitor the performance of the health care delivery system. For example, these studies address health care reform issues. Including the spiraling costs of health care and the increasing number of poor, uninsured, and other persons with inadequate access to basic health care services. Topics examined by AHCPR grantees include:

- Primary care, the site of most "first contact" care and the source of most referrals to secondary and tertiary care.
- Market forces, including cost containment efforts and their effects on the supply, quality, and costs of health services.
- Managed care and the assessment of the quality of the processes and outcomes of coordinated care.
- Costs of treating AIDS and improving treatment for persons with HIV.
- Rural health care, which differs significantly from urban health care.
- Infant mortality, reducing its incidence and improving access to cost-effective preventive services.
- Medical liability and malpractice reform to reduce associated health care costs.
- Health care of the aged and disabled, including long-term care.

Policy studies conducted by AHCPR analyze how Americans use and pay for health care. AHCPR also examines personal health insurance coverage; the quality, supply, and costs of hospital services; and long-term care service needs and utilization.

AHCPR conducts and supports evaluations of the risks, benefits, effectiveness, and where possible, cost-effectiveness of devices, procedures, and other medical technologies used in preventing. diagnosing, treating, or managing certain conditions.

Ways to improve the effectiveness and appropriateness of clinical practice

The following activities are components of AHCPR's Medical Treatment Effectiveness Program (MEDTEP):

Patient outcomes research. "Outcomes"—the effects of patient care—include traditional measures of morbidity and mortality, as well as quality-of-life factors, such as how patients feel after treatment and how readily they resume daily activities. Outcomes research tells us what works best, for whom, and at what cost. These studies compare the outcomes of different ways of preventing, diagnosing, treating, and/or managing a particular condition. The largest AHCPR-funded outcome studies are Patient Outcomes Research Team (PORT) projects: 5-year studies conducted by experts from multiple clinical and scientific fields. Among the PORTS supported by AHCPR are studies on diabetes, pneumonia, heart attack, stroke prevention, prostate disorders, low birthweight, and back pain. AHCPR also supports other MEDTEP research projects, including the study of outcomes associated with pharmaceutical therapy, analysis of minority health care issues, and the development of outcomes research methods.

Clinical Practice Guidelines. AHCPR sponsors multidisciplinary panels of private-sector health care experts and consumers, and contracts with private nonprofit organizations to develop clinical practice guidelines. Guidelines help health providers make better medical care decisions, reducing ineffective or inappropriate services.

U.S. DEPARTMENT OF HEALTH
AND HUMAN SERVICES
Public Health Service

Guidelines result from extensive literature reviews and reflect the best scientific evidence. They are peer-reviewed and tested through onsite clinical evaluations by potential users to assess their validity, efficacy, and applicability. AHCPR-supported practice guideline topics include management of acute pain; urinary incontinence in adults; prediction and prevention of pressure ulcers; management of functional impairment due to cataract; sickle cell disease in newborns and infants; diagnosis and treatment of benign prostatic hyperplasia; diagnosis and treatment of depression; HIV/AIDS; mammography; and others.

Data bases for research. Health services researchers require large amounts of data to study health care as it is delivered to the general population. AHCPR is working to create more uniformity among health care data bases, including assessment of the utility of automated medical records, and supports the development of improved research data bases.

Dissemination. AHCPR administers an extensive dissemination program to ensure widespread availability of research findings and clinical practice guidelines. AHCPR also funds studies that contribute to understanding effective dissemination.

AHCPR Operation and Organization

AHCPR responsibilities are carried out through the following components:

Center for General Health Services Extramural Research. Awards grants and contracts to study the ways health care is organized, delivered, and financed. Focuses basic research on quality of care, health status measurement, primary care, cost-effectiveness of interventions, and access to appropriate care. Conducts the AIDS Cost and Service Utilization Survey: a national examination of how persons with AIDS use and pay for health and social services.

Center for General Health Services Intermural Research. Through large data bases, analyzes health insurance coverage, access to health care by groups of special policy concern, and national and regional health care costs and expenditures; examines hospital costs and services and long-term care issues.

Center for Medical Effectiveness Research. Awards grants and contracts for patient outcomes research and studies on practice variations, outcomes research methodologies, and other medical effectiveness research topics. Monitors the work of Patient outcomes Research Teams and the MEDTEP Research Centers on Minority Populations.

Center for Research Dissemination and Liaison. Publishes AHCPR reports and disseminates information to a wide array of audiences; collaborates with with National Library of Medicine to make information more accessible. Awards grants to study the dissemination, use, and effects of health services research findings and clinical practice guidelines. Supports training programs and conferences for health services researchers.

Office of the Forum for Quality and Effectiveness in Health Care. Arranges for the development, periodic review, and updating of clinical practice guidelines. Supports development of performance measures, standards of quality, and medical review criteria for use by health care practitioners and others in reviewing health care quality and services.

Office of Health Technology Assessment. Evaluates the safety, efficacy, effectiveness, and where possible, cost-effectiveness of health care technologies used to prevent, diagnose, treat, or manage certain conditions. These evaluations often form the basis for coverage decisions by federally funded medical programs.

Office of Science and Data Development. Supports and conducts activities designed to increase the amount and usefulness of data (such as that from health insurance claims data bases and computer based patient research) for outcomes and other health services research.

Office of Program Development. Coordinates the development of the research and policy programs and priorities of AHCPR. Develops and analyzes legislative proposals. Manages the activities of the National Advisory Council for Health Care Policy, Research, and Evaluation.

Office of Scientific Review. Responsible for conducting and reporting the results of the peer review of research grant applications and contract proposals.

Office of Management. Provides overall management of the AHCPR grant and contract programs. Formulates and executes the AHCPR budget. Provides administrative and management support services to AHCPR.

For more information, contact:
Division of Communications
Agency for Health Care Policy and Research
Executive office Center, Suite 501
2101 East Jefferson Street
Rockville, MD 20852
Telephone: (301) 594-1364
Media inquiries only: (301) 594-1370
TDD (Telecommunication Device for the Deaf): (301)594-6173

AHCPR Pub. No.93-0055

AHCPR

fact sheet

Agency for Health Care Policy and Research

January 1992

2101 East Jefferson Street, Suite 501, Rockville, MD 20852, 301/227-8364

AHCPR-Commissioned Clinical Practice Guidelines

Clinical practice guidelines are systematically developed statements designed to help physicians and their patients in making decisions about appropriate care for specific health conditions. Guidelines reflect current scientific knowledge of practices and the best professional judgment used to prevent, diagnose, treat, or manage diseases and disorders.

Why Develop Guidelines?

Research during the past two decades has identified major variations in the way physicians care for a specific health problem. Doctors in neighboring communities, for example, have been found to differ dramatically in their decisions about performing hysterectomies, prostatectomies, and tonsillectomies. Although some differences in medical care are expected, significant, unexplained variations raise concern that some patients are not being treated effectively and, in some cases, may be harmed. There also is concern that health dollars spent by patients or their families, insurers, employers, and Federal and State programs such as Medicare and Medicaid are being wasted on ineffective or inappropriate care.

Researchers believe that practice variations occur in part because there is no strong consensus among physicians about what works best and for whom. Faced with often contradictory reports in the health care literature, clinicians rely on their own professional judgment, or the judgment of their peers, when making practice decisions.

In 1989, Congress established the Agency for Health Care Policy and Research (AHCPR) to improve the quality, appropriateness, and effectiveness of health care in the United States. As part of its congressional mandate, AHCPR facilitates the development of clinical practice guidelines by commissioning expert panels to address selected clinical conditions.

AHCPR-commissioned guidelines recommend ways of diagnosing, treating, managing, and where possible, preventing conditions. The guidelines are based on a comprehensive review of the scientific literature, on valid evidence presented at open meetings, and on the professional judgments of panel members and other experts in the field. The guidelines also reflect how health care practices affect patients (these effects are known as "outcomes") and take costs of care into account. In contrast to other guidelines, those developed under AHCPR sponsorship are the product of many different types of health providers (primary care physicians and specialists, nurses, and allied health care providers such as physical therapists) and consumer representatives. Some panels include experts in other areas as well, such as psychology and ethics. Panels supported by AHCPR and organizations under contract to the Agency are currently developing guidelines for conditions or procedures such as congestive heart failure, low back problems, otitis media (ear infection) in children, urinary incontinence in adults, stroke rehabilitation, depression, and the requirements for high quality mammography.

AHCPR believes that with widespread use of clinical practice guidelines:

- Doctors, nurses, and other health care providers will make practice decisions that rely more on science-based knowledge and respected professional judgment than is currently the case.
- Patients will become better informed health care consumers who can work as partners with their health care providers.
- Unnecessary health care practices will be eliminated or greatly reduced.
- The quality of health care will improve.
- The overall costs of health care may decrease.

How Are Guidelines Developed?

AHCPR's Office of the Forum for Quality and Effectiveness in Health Care oversees the development of clinical practice guidelines. The

BEST COPY AVAILABLE

Office convenes panels of physicians, nurses, other health care experts, and consumer representatives to develop a guideline, or contracts with public and private nonprofit organizations to do so.

Broad collaboration is critical to the success of the guidelines, which depend on voluntary adoption by practitioners. The process benefits greatly from the full participation of medical specialty societies, health care organizations, academic medical centers, scientific bodies, and organizations that set standards, measure quality, and conduct research. Although it is not possible for all of these groups to directly participate in each panel, they play an important role in the peer review and later the field review of the guidelines.

Selection of Guideline Topics

AHCPR considers a number of factors in selecting guideline topics, including:

- Adequacy of science-based evidence.
- Number of individuals affected by the clinical condition.
- Expected potential for reducing inappropriate variations in the prevention, diagnosis, management, or outcomes associated with the condition.
- Specific needs of Medicare and Medicaid populations.
- Cost of the condition to all payers, including patients.

Selection of Panel Members

AHCPR requests nominations for panel members from a variety of interested parties, including practitioner and consumer organizations. Panels range from 9 to 15 members, and expert consultants are used as needed. AHCPR announces in the *Federal Register* that it, or its contractors, are establishing new guideline panels. The announcements call for nominations by a certain date and list the selection criteria. The panel chair or co-chairs (some panels have two) must have the following qualifications: training and clinical experience relevant to the condition being examined; demonstrated research and leadership on the condition, including publication of relevant, peer-reviewed articles; a broad public health perspective on the condition or procedure being addressed; demonstrated ability to both lead a health care team in the decisionmaking process and in responding to consumer concerns; and prior experience in developing guidelines. The criteria for panel members are similar.

AHCPR's Administrator selects the chair or chairs and the panel members. The Administrator then appoints the selected nominees.

Methodology Used to Develop Guidelines

In developing a guideline, a panel considers the scientific evidence, weighs the consequences of different options, and explicitly describes how final recommendations were made. Four basic guideline development steps are carried out:

1. *Definition of the Process.* The panel's first task is to define the goals and scope of the project. This includes clarifying the target condition; identifying the providers, practice settings, and types of patients for which the guideline is intended; determining the clinical interventions that will be considered; specifying the desirable and undesirable patient outcomes (clinical benefits and harms) that can result from these interventions; and selecting the types of scientific evidence that need to be examined.

To organize the task of determining how the clinical outcomes will be influenced by various interventions, the panel develops a flowchart. This tool serves as a "road map" for the review of the scientific evidence and the assessment of the clinical benefits and harms.

2. *Assessment of Clinical Benefits and Harms.* The panel assesses relevant data from the available scientific evidence to evaluate the clinical benefits and harms of the various interventions. A consultant skilled in determining the specific validity of various clinical studies prepares a literature review for the panel members. In some situations, an evaluation technique called "meta-analysis" may be used. Meta analysis, which combines or synthesizes the results of related scientific studies, permits conclusions to be drawn from studies with sample sizes that otherwise would be too small to be statistically useful. These findings are combined with expert opinions to develop an "appropriateness profile"— a listing of appropriate, inappropriate, and "gray zone" interventions that represents an early version of the eventual practice guideline.

3. *Assessment of Health Policy Issues.* Panel members determine how they will address health policy issues such as costs, medical liability, use of limited resources, and concerns of patients, practitioners, and payers that may interfere with implementation of the guideline. After the panel prepares and reviews a written report on these issues, an open forum is announced in the *Federal Register*. During the forum, individuals and group representatives present testimony on health policy issues of concern to them. A panel session follows, during which members determine how to address the expressed views.

4. *Preparation of Practice Guideline Document.* A draft version of the guideline is prepared and reviewed informally by panel members. The draft is then revised and distributed to outside experts for peer review. Comments from peer reviewers are circulated among panel members, and the draft guideline is further revised. Then a cross-section of intended users and consumers field review the guideline to determine how

readily it can be adopted in clinical practice settings. Findings from these field reviews are summarized and distributed to panel members.

The last step is systematic evaluation of the comments received through peer review and field review. Panel members agree on the criteria to be used in addressing comments and results. Then final documents are prepared.

How Are Guidelines Made Available?

Each guideline is prepared in several formats for use by consumers, health care practitioners, the health care industry, policymakers, and researchers in health services and biomedicine. A long, technical version, the Guideline Report, provides scientific evidence to substantiate the recommendations and findings in the guideline. A shorter version, the Clinical Practice Guideline, can be used by practitioners as a "shelf reference"; and an even more abbreviated version, the Quick Reference Guide for Physicians, can be used as a "desk reference." Additionally, the panels prepare a plain-language Patient's Guide in English and Spanish. Guideline information also will be available on-line through the National Library of Medicine's computer services, medical libraries, and indexing services.

As guidelines are completed, AHCPR promotes wide dissemination to practitioners and other users. AHCPR encourages health care professional organizations, hospital associations, academic medical centers, medical educators, third-party payers, medical care reviewers, and others to distribute the guidelines to their members and constituents.

Guidelines will be reviewed and updated as necessary. Many patient outcomes and effectiveness studies are now in progress, including Patient Outcomes Research Team projects and smaller-scale studies supported by AHCPR. Findings from these studies will contribute to the development and updating of current and future clinical practice guidelines.

In addition, AHCPR will evaluate the adoption of the guidelines and their effect on health care practice.

How Can the Public Participate in Guideline Development?

AHCPR issues requests for public comments in the Federal Register as each new guideline panel begins its work. The panels are particularly interested in the following information: scientific evidence and studies related to the prevention, diagnosis, treatment, and management of the medical conditions; outcomes of importance to patients; risks and benefits; geographic variations in practice; cost-effectiveness; and potential effects on access to care.

Individuals and organizations with information on specific guideline topics should submit them in writing to:

Kathleen McCormick, Ph.D., R.N.,
Director
Office of the Forum for Quality and Effectiveness in Health Care
Agency for Health Care Policy and Research
Executive Office Center, Suite 401
2101 East Jefferson Street
Rockville, MD 20852

TOTAL P.04

Discussion

on the

Use of Consultation and Consensus-Building Processes

for Implementing the Clean Air Act of 1990

Office of Air And Radiation

U.S. Environmental Protection Agency

Revision Date: March 4, 1992

I. Introduction

The purpose of the enclosed materials is to describe several alternative approaches for consulting with external interests as EPA develops regulations that implement the Clean Air Act. All approaches assume that the overall goal of the rulemaking process is to produce rules that are not only technically sound, fair, and effective, but have the greatest likelihood of being legally defensible and implementable (thereby achieving environmental results in a timely and efficient manner). In order to achieve this goal, EPA should ensure that three types of concerns are addressed by appropriate members of the Agency in conjunction with each rulemaking effort:

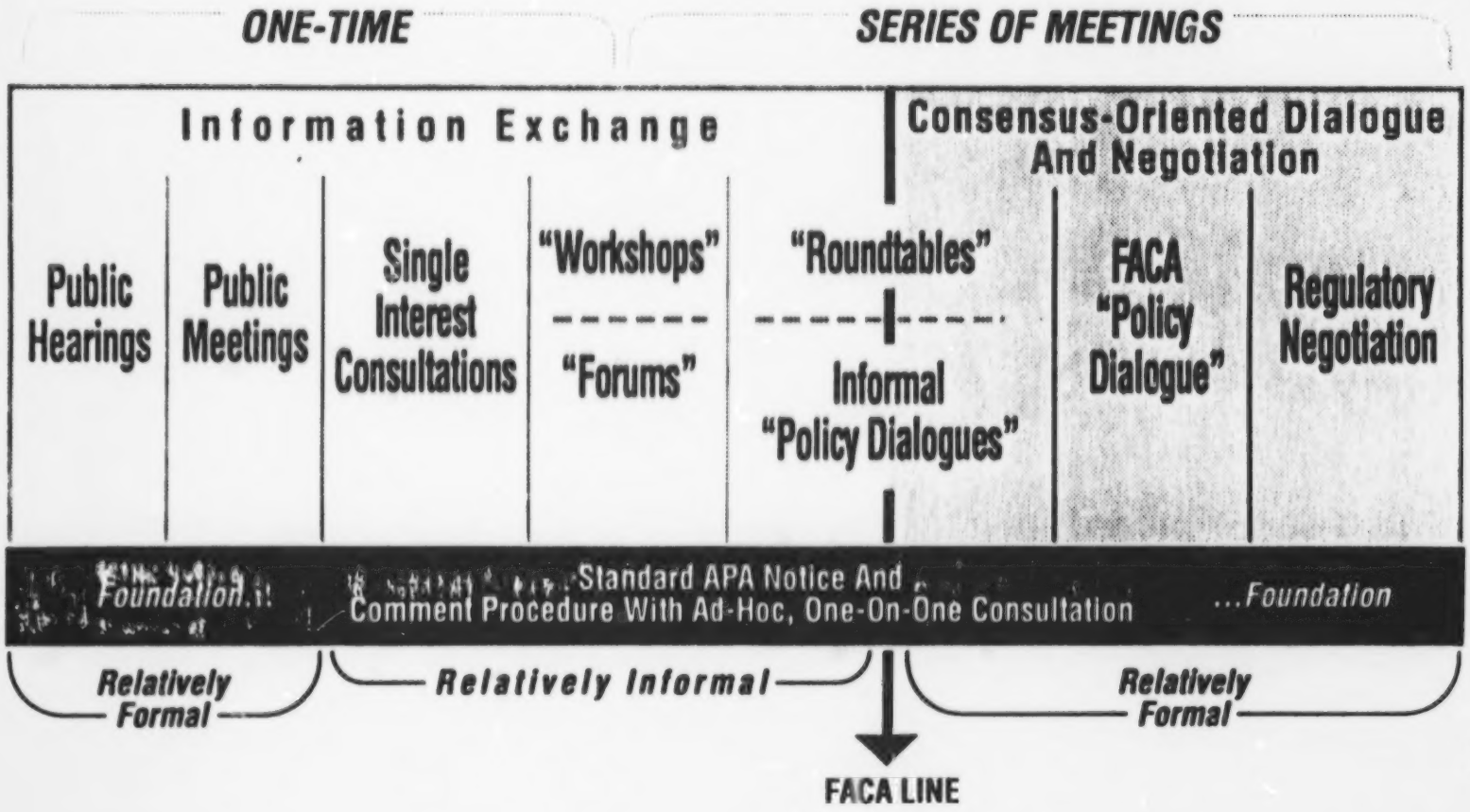
- **Technical:** What are the relevant technical issues of concern and how can they be resolved?
- **Procedural:** What potential procedural issues may interfere with implementation of the rule?
- **Political:** Who has the power to influence or block implementation of a proposed rule? How have their concerns been responded to, accounted for, or addressed?

If, at the close of a rulemaking process, significant issues of concern to major interest groups remain in any of these three categories, it is more likely that the rule will not be implemented on time. Therefore, these questions should be carefully considered by each member of the EPA rulemaking work group and other appropriate EPA staff, both at the outset and at critical junctures in a specific rulemaking effort.

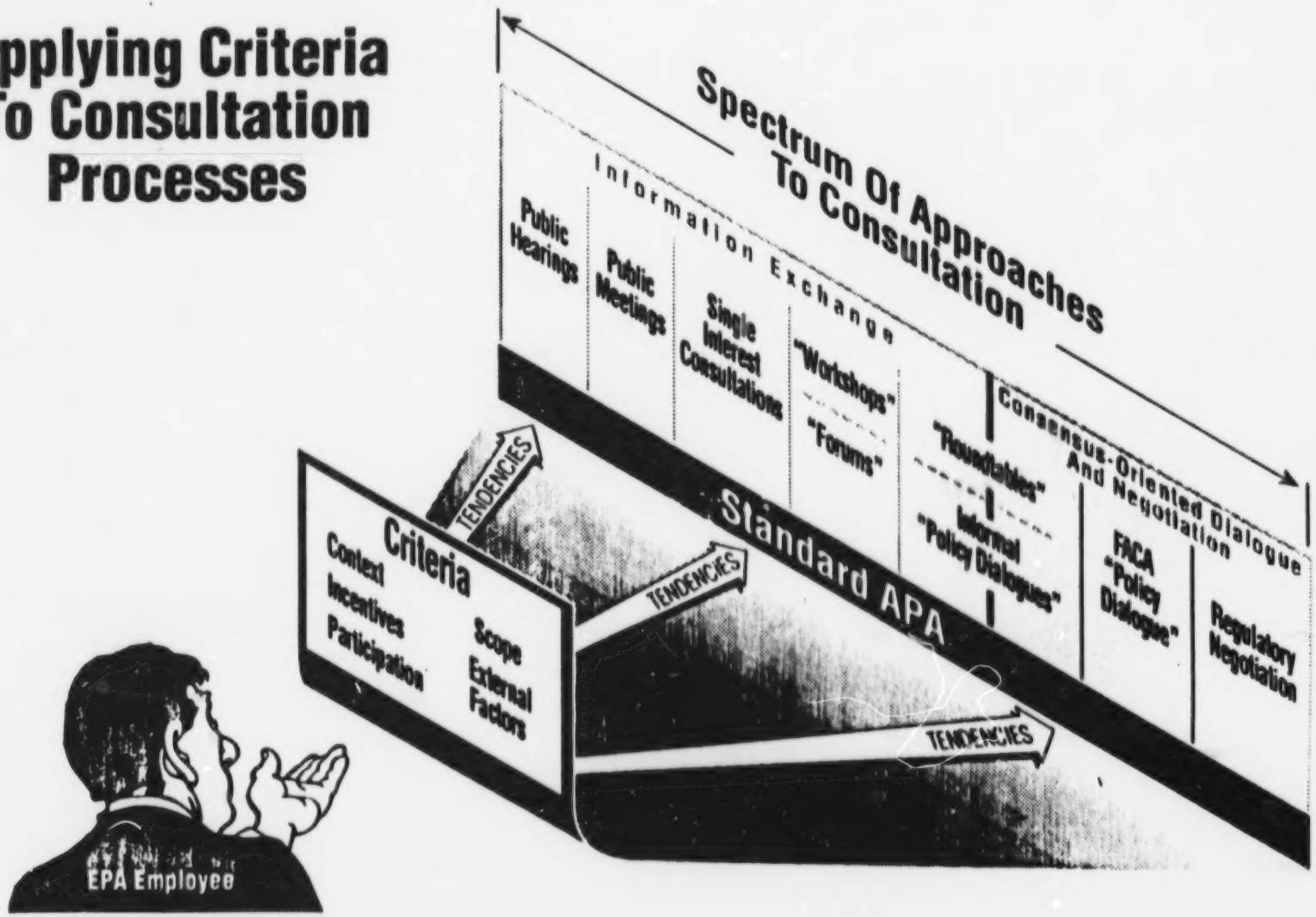
Assuring that there are adequate answers to all of the questions posed will not "guarantee" a positive outcome. There are potential negative consequences to attempting some of the processes. For example, some of these processes can consume agency time and resources. They can create additional expectations from participants (e.g. that EPA will incorporate all of their comments into a proposal). However, answers to the questions above should inform the design and outcome of the rulemaking process whether a consultation process is pursued or not.

In Part II below, a description of five parts of the consultation process spectrum is provided. In Part III, criteria for selecting an appropriate process are provided. Part IV presents tendencies in the application of criteria to the choice of consultation processes. Part V provides a checklist of key factors to be considered before initiating any consultation process. The graphics which follow portray how the criteria and tendencies described in Parts III and IV can be used by EPA staff to determine which consultation and/or consensus-building process may be appropriate for their circumstances.

Spectrum Of Consultation And Consensus-Building Approaches



Applying Criteria To Consultation Processes



II. Parameters of Six Consultation Processes for Implementing the Clean Air Act

The following discussion outlines the parameters of six consultation processes, in addition to the activities conducted under the Administrative Procedures Act. Each of these processes should be viewed as a range of activities, or part of a "spectrum," which may overlap with one another. This "spectrum" of approaches varies according to three general characteristics: 1) the degree to which each process involves interactions between EPA and each interest group separately or several interests groups at the same time; 2) whether the process is a one-time event or a series of discussions and interactions; and 3) the degree to which the objective of the process is to exchange information and perspectives or to develop consensus solutions.

EPA, as with all federal agencies, is subject to and must comply with the requirements of the Administrative Procedures Act (APA) and the Federal Advisory Committee Act (FACA). A brief description of APA activities is presented below for information purposes. Within the spectrum of additional consultation processes, the first approach described is "single interest group consultations," which involves meetings with representatives of a single interest group. On the other extreme of the spectrum, "regulatory negotiations" involve all affected interests in a series of formal negotiation sessions aimed at achieving consensus agreements. The middle range includes one-time or a series of interactions that have the objective of information exchange, fact-finding or consensus-building in relatively less formal settings. When EPA seeks consensual advice from select groups of experts, EPA will follow FACA the requirements that the Agency obtain a charter, have balanced membership, and announce and conduct the meetings. Each process has general characteristics, but should not be viewed as completely separate and distinct. Rather, activities could be contemplated that would fall both within and between some of the methods described below.

Administrative Procedures Act (APA)/Public Hearings

What: Information sharing activities such as publication of significant decisions in the Federal Register and conduct of hearings to obtain public comment prior to specific decision points.

Objective: To provide an opportunity for all affected interests to formally submit comments on proposed rules.

Who: Public at-large and representatives of key interest groups.

When: At specific decision points, as required by APA.

Role of EPA: Rule proponent, decision maker.

Example: Any formal public hearing or public comment period that precedes or follows issuance of a proposed rule.

Public Meetings

What: Information sharing activities such as meetings conducted prior to public hearings. Also includes contacts with and inquiries from key affected interests, either in small meetings or by telephone.

Objective: To provide an opportunity for all affected interests to ask questions and share concerns. Allows exchange of information with interested and affected parties in less formal, one-on-one settings.

Who: Public at-large and representatives of key interest groups.

When: Usually prior to formal decision points of the regulatory process.

Role of EPA: Rule proponent, decision maker.

Example: Any informal public gathering in which the Agency imparts information and receives comments reflecting a lack of consensus on an upcoming regulatory or policy action. Any telephone or personal consultation with individuals,, or groups not established by the Agency.

Single Interest Group Consultations

What: One time or a series of meetings, briefings or consultations conducted separately with individual representatives of distinct interest groups.

Objective: To give and receive information with representatives of a single interest group; to exchange views, utilize the expertise of external constituencies, build trust and improve relationships.

Who: Selected individual representatives or self-selected delegations of each major affected interest.

When: At beginning of and throughout the regulatory process, as needed.

Role of EPA: Principal convenor and participant. EPA retains policy/decision making role.

Example: EPA's Petroleum Refinery Cluster information gathering meetings with industry, states, and environmental/citizen/labor groups.

Information Exchange "Forums"/"Workshops"

What: One-time events (one to three days in length) involving diverse and balanced representation of all affected interest groups.

Objective: To communicate to interest groups the issues, concerns, and constraints facing the agency in the rulemaking; offer an opportunity for informal input by each individual interest group; allow each individual interest group to hear the perspectives and concerns of other interests; and to allow each individual interest group to hear EPA's responses to inquiries from those present. EPA does not seek to obtain consensus.

Who: All relevant interest groups, designated or self-selected, but limited representatives for each. Balanced representation.

When: Typically, early in the regulatory development process.

Role of EPA: Participant and in some instances principal convenor. EPA retains policy/decision making role.

Example: The Keystone Center's RCRA Reauthorization Forum

"Roundtables"/Informal "Policy Dialogues"

What: Ongoing, informal information discussions over regulatory policy issues with designated representatives of affected interests. EPA neither seeks nor obtains consensus advice.

Objective: To give each individual interest group the opportunity to gain information, discuss issues, and generate options and alternatives. Other objectives of single interest group consultations and information exchange forums may also apply.

Who: All relevant interest groups, designated or self-selected, but limited representatives for each. Balanced representation.

When: At beginning of and during the regulatory decision making process, as appropriate.

EPA Role: Participant and typically the principal convenor. EPA can observe or participate in policy dialogues that are convened by outside parties. EPA may also use advice and recommendations generated by committees established by outside parties (In such cases, FACA may apply, warranting consultation with the EPA FACA attorney in Office of General Counsel).

Example: Clean Air Act Permits Roundtable, CAA Voluntary Reductions Roundtable.

FACA "Policy Dialogue"

What: Formal "policy dialogue" that follows the rules of the Federal Advisory Committee Act. EPA's goal on establishing such a committee is to reach consensus.

Objective: In a more formal setting than the above options, obtain advice, sometimes in the form of consensus recommendations, on significant policy issues through a public meeting process that involves representatives of a balanced group of affected interests. Objectives of other processes may apply.

Who: Designated representatives of all relevant interest groups. Representatives may nominate themselves, for each. Balanced representation.

When: At beginning of and at regularly scheduled intervals throughout the regulatory process.

EPA Role: EPA charters the FACA committee and makes sure all other FACA requirements are met. EPA staff may be a committee member or may participate in discussions at the request of the committee. If a consensus is achieved EPA should give the highest possible consideration to recommendations.

Example: Acid Rain Advisory Committee.

Regulatory Negotiations

What: Formal negotiations under the Negotiated Rulemaking Act and FACA between designated representatives of affected interests aimed at achieving consensus agreements on proposed rules.

Objective: To negotiate consensus agreements on the content of and/or principles to be reflected in a particular rule. The agency participates as a party to these negotiations and commits to using any consensus recommendations that may be reached as the basis of a proposed rule. Other objectives of other consultation processes also may apply.

Who: All relevant interest groups with designated representatives for each. Balanced representation.

When: At beginning of and at regularly scheduled intervals throughout the regulatory process.

EPA Role: Party to the negotiations. EPA retains ultimate policy/decisionmaking role. As per requirements of the APA, EPA must be responsive to comments received on whatever proposed rule emerges from the "reg neg" process.

Example: Reformulated/Oxygenated Fuels and Fugitive Emissions Regulatory Negotiations.

III. Assessment Criteria for Selecting Consultation Processes

Selecting the appropriate consultation method is a critical step toward a successful rulemaking effort. The people involved, the broader ramifications of the rule subject matter, the technical and political complexity of the effort, the inherent incentives for participation by all interested parties, as well as timing and resources, all play a key role in choosing the most appropriate consultation method. Below are criteria that managers and work group planners should consider in developing a strategy for carrying out their charge.

Context

- What is the larger context for the rulemaking? For example, what other rules may be under discussion with the interested parties?
- How might these or recently completed rulemaking processes effect the viability of using a consultation process in conjunction with the current rule?
- What other goals is the agency trying to accomplish in the given timeframe? Are the goals consistent with each other? Are the potential outcomes of a consultation process consistent with agency goals?
- Is a consultation process appropriate given the current circumstances (e.g., if parties are likely to litigate regardless of the result of a consultation process for the rulemaking, it may not be in EPA's interest to use a consultation process to achieve its desired outcome).
- Is EPA internally organized such that it will be able to participate and be perceived as a unified team?

Scope of Issues

- What is the scope of issues that could be addressed in a consultation process? Are issues too numerous or too complex to address with multiple parties? If not, how can the number and complexity of issues be most effectively managed?
- Is technical research or information available that can resolve some questions and narrow the scope of issues of concern?
- Do the issues of concern for various parties overlap? Given the scope of issues, what consultation process is appropriate to address them? For example, some parties may be concerned about certain aspects of a rule while other parties have concerns about separate aspects. In this case, single interest group meetings may be an appropriate first step in addressing parties' concerns.

Participation

- Who is potentially affected by the regulation and who has the power to influence or block the implementation of the regulation -- and therefore may be an appropriate candidate for participation in a consultation process?

- o For multi-party processes, can representatives or spokespersons of all relevant interest groups be identified?

- o Can interest groups agree on individuals that can represent them?

- o Is the representation of different interest groups balanced?

- o Are the number of people contemplated a reasonable number for the process under consideration?

- o Does the Agency know of any interest groups that may not be aware of a particular rulemaking effort or that are aware but do not intend to participate? If so, should their participation be encouraged by the Agency?

Incentives

- o What are the agency's incentives for conducting a consultation process (i.e. what might the agency gain)?
- o What is the agency's best alternative to using a consultation process to develop legally defensible and implementable rules? Are the alternatives potentially more beneficial or less risky than the proposed process?
- o What are the parties' incentives to participate in a cooperative effort with EPA? What incentives or disincentives does EPA create by the choice it makes in suggesting one consultation approach versus another? (i.e. Do incentives for parties to participate vary if EPA is participant rather than an observer? If the meetings are information exchange versus consensus-oriented?)

- o What are the parties' underlying interests? Are they likely to be advanced through participation in the process?

- o What are the parties' best alternatives to participation? Are alternatives likely to be perceived as more beneficial or less risky than the proposed process?

Timing

- o How much time is available to prepare for and conduct a consultation process? Is that time adequate for the type of process contemplated? For example, if the timeline for a rule is constrained, it may not be appropriate to enter into an ongoing consultation process such as a FACA "policy dialogue" or regulatory negotiation.

- o Is the timeframe of the consultation process in conflict with other relevant

activities of the parties' or the agency? For example, will the demands of other responsibilities afford adequate staff time for the agency to pursue a consultation process?

Resources

- What kind of resources (people, materials, and financial) are necessary to conduct the consultation process? Are adequate resources available to conduct the process?

IV. Applying Criteria to Consultation Processes

Combined responses to the questions presented above indicate tendencies for the most appropriate consultation approach to use for increasing participation in a regulatory effort. This section offers guidance in applying the criteria discussed in Part III to the consultation processes presented in Part II.

The first consideration for the agency is whether to conduct any type of consultation process above the requirements of the Administrative Procedures Act (APA). Agency staff should begin by considering the answers to the contextual and incentive questions posed in Part III. If there appear to be conflicts between rulemaking processes or if parties are likely to litigate regardless, it may be appropriate to delay or avoid meetings, workshops or dialogues. Similarly, if the agency's or parties' incentives to participate appear to be inadequate, the agency may decide not to conduct any type of process beyond APA requirements.

Potential reasons to conduct a consultation process include the following:

- Obtaining information from external interests to improve the content and implementability of a rule or policy;
- Improving relationships with affected interests; and
- Creating new ideas or alternative policy or regulatory structures.

If the agency decides to proceed with a consultation process, several decisions will assist the agency in narrowing the choice of a specific consultation process. These decisions arise out of answering the following five questions:

- Should the agency pursue information exchange meetings or consensus-oriented dialogues or negotiations?

- o Should the agency conduct a one-time meeting or a series of meetings?

- o Should the agency meet with one interest group at a time or with all interest groups simultaneously?

- o Is the agency interested in conducting relatively informal or more formal meetings?

- o If the agency decides to pursue formal meetings, should it conduct a FACA policy dialogue or regulatory negotiation?

Before deciding on any type of consultation process, the agency should determine that it has adequate time and resources to conduct the process.

Should the agency pursue information exchange meetings or consensus-oriented dialogues or negotiations?

In deciding between information exchange meetings or consensus-oriented meetings, the agency should consider answers to the questions of Scope and Participation presented in Part III.
Information exchange meetings are more suitable if:

- o The scope of the issues is ill-defined or too numerous or complex to address in dialogue setting;

- o Interested parties are not sufficiently educated on the issues to participate in a decision-oriented process;

- o Individuals or groups interested in the issues cannot be easily identified or are poorly organized;

- o Interest groups cannot support selection of representatives to participate in a multi-party process.

By conducting information exchange meetings, interest groups may be able to organize themselves and select representatives to later participate in a consensus-oriented process. Similarly, if affected interests are inadequately informed about relevant issues, conducting information exchange processes may prepare them to participate effectively in dialogues or negotiations.

There may be interests who should be aware of and involved in a regulatory or policy decision, but have not been involved to date. In this case, the agency may identify individuals to participate in information exchange meetings in order to bring them up to speed for effective involvement.

In order to conduct consensus-oriented dialogues or negotiations, representatives of affected interests must be identifiable. Parties must be able to support the selection of representatives to speak for them during discussions. Issues must be somewhat focused and discreet in order to effectively discuss them (though policy dialogues often address broad sets of issues). Parties should be relatively knowledgeable about the issues before convening them for consensus-oriented discussions.

In many cases, information exchange meetings will meet the goals of the agency. Instances in which the agency may want to consider consensus-oriented processes include situations where:

- o Parties seek a consensus effort and are willing to commit the time and resources necessary to participate;
- o Agency is willing to allow consensus decisions to guide a policy or regulatory decision;
- o Agency desires an outcome supported by all affected interests;
- o Agency seeks a policy or regulatory outcome that is more easily implemented (i.e. reduced administrative and legal challenges);
- o Agency needs to address competing points of view earlier rather than later in the process; or
- o Agency perceives that the incentives, timing, and available resources are in place so that the process is likely to succeed.

Should the agency conduct a one-time meeting or a series of meetings?

Due to the nature of consensus-oriented meetings, these processes usually involve more than one meeting and are often subject to FACA. In the case of an information exchange consultation process, the agency may conduct a one-time meeting or a series of meetings. Primarily, this decision is determined by whether a one time meeting can provide sufficient technical and political information to complete the regulatory or policy task. If one meeting suffices, the choice is obvious. If affected interests may view one meeting as inadequate to communicate their concerns, more than one meeting should be considered.

For more than one meeting, time constraints and availability of resources will affect the number of meetings conducted. The agency should clearly identify the scope of issues, resources available, and the goal of conducting meetings in order to determine in advance the number and timing of meetings. The agency should indicate its intentions for the number of meetings from the outset. This indication sets reasonable expectations and

allows affected interests to plan their participation.

Should the agency meet with one interest group at a time or with all interest groups simultaneously?

All consensus-oriented meetings, by definition, involve more than one interest group. If the agency proceeds with an information exchange consultation process, it may meet with interest groups together or separately. If parties are interested in meeting together, it may be prudent and responsive to conduct public hearings, public meetings, workshops or roundtables. The agency should meet with parties simultaneously only if they are willing to share their views in each other's presence. The agency may want to have parties meet together if it wishes for parties to hear each other's views. If the nature of parties' concerns overlap, it also may be more efficient to have parties meet together so that similar concerns can be addressed at once.

Is the agency interested in conducting relatively informal or more formal meetings?

For information exchange processes, public hearings and meetings are considered more formal processes, whereas workshops, forums, and "roundtables" are less formal in nature. If the agency wishes to create an opportunity for a formal public testimony record, it should conduct a public hearing. A public meeting offers more opportunity for a two-way exchange of information than a hearing. Workshops, forums, and "roundtables" are sometimes off the record (though the press may be in attendance) and more discussion--rather than question/answer--oriented than public meetings.

For a processes that aim at generating a consensus agreement, the agency must establish a FACA "policy dialogue" NOTE: Not all FACA efforts need be consensus-oriented.) If the meetings, issues and proposed participants meet FACA triggers, a FACA process must be pursued. If the agency seeks a high degree of formality in order to legitimize their decision making process, a FACA "policy dialogue" or regulatory negotiation may be appropriate. If the agency seeks to involve participants of high stature in the process, the formality of a FACA "policy dialogue" or regulatory negotiation may be appropriate.

If the agency decides to pursue formal meetings, should it conduct a FACA policy dialogue or regulatory negotiation?

If the agency seeks the involvement of a consistent group of people in a formal setting, it must decide whether a "policy dialogue" or regulatory negotiation is more appropriate. The distinction here is relatively simple. If the scope of issues suggests broader policy issues, a policy dialogue is appropriate. Consensus agreements for a specific rulemaking are sought through the regulatory negotiation process.

The tendencies toward use of particular consultation processes presented here are

a sample of how the criteria (Part III) can be applied to the consultation processes (Part II). Each situation should be carefully evaluated using each of the criteria before deciding on a specific approach.

V. Key Factors in Initiating Consultation Processes

Using the criteria and tendencies described in Parts III and IV above, the agency staff should begin its particular Clean Air Act implementation task by assessing the regulatory situation and making some preliminary judgements about what, if any, consultation process may be appropriate. The Agency staff should consider the following checklist of internal and external factors as they proceed.

Internal

___ Does the Agency have adequate time and resources to conduct the consultation process under consideration?

___ Does the Agency have adequate incentive to conduct the consultation process?

___ Has the Agency decided its role in the process and remained cognizant of the affect of its role on the interest group's incentives to participate?

___ Has an appropriate scope of issues been identified?

___ Are other processes being conducted on related issues (either within the office or in other offices)?

___ Will other processes affect or be affected by the consultation process? Will any related efforts be coordinated with the proposed process?

___ Have the appropriate roles within EPA been defined for the effort? Has headquarters been consulted sufficiently early about the selected consultation process and the design of it? Is the proposed EPA team effective? Has an appropriate individual been identified to lead the effort?

___ Are appropriate on-going channels of communication with headquarters in place?

External

___ Does the Agency have an overall strategy or plan for consulting with outside groups on the particular issues of concern?

___ Does the Agency know one or two key representatives in each interest group in order to conduct an initial assessment?

___ Have the Agency's expectations of the outside interests been defined and communicated to the interests?

___ Does the Agency know the major concerns or interests of each interest group?

___ Does the Agency know the current activities of affected interest groups?

___ Is the Agency aware of the key interest group's alternative means of effecting the regulatory process?

___ Does the Agency have reliable information and feedback from affected interests as to the viability of the selected approach?

___ Is the Agency likely to have the participation of the key interest groups in a consultation process?

___ Does the Agency have an understanding for the needs and concerns of agencies and government officials outside EPA (e.g. White House, states, congressional members)?

___ Does the Agency have a plan for coordinating the input of these outside non-government offices?

In conducting the assessment or as part of conducting a consultation process, the Agency may want to use the services of a third party neutral. Possible roles for third parties may include:

-- providing EPA staff with feedback on their preliminary judgements about what consultation strategy to pursue;

-- exploring and obtaining feedback from affected interests on EPA's ideas about strategies to pursue;

-- if necessary, conducting a formal "convening assessment;" and

-- facilitating discussions/negotiations with key interest groups.

Within the Office of Policy, Planning and Evaluation (OPPE), there exists the Regulatory Negotiation Project. The staff affiliated with this project can be of assistance to Agency staff as they pursue not only regulatory negotiations, but all consultation and consensus-building options discussed in this paper. The Agency's Office of Policy, Planning and Evaluation staff, EPA,s Committee Management Officer, and the Office of General Counsel FACA attorney can be of particular assistance in answering questions related to FACA and the use of outside third party facilitators.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

FACT SHEET

NEGOTIATED RULEMAKING / REGULATORY NEGOTIATION

WHAT IS NEGOTIATED RULEMAKING?

Negotiated rulemaking is a process which brings together representatives of various interest groups and a federal agency to negotiate the text of a proposed rule. The goal of a negotiated rulemaking proceeding is for the committee to reach consensus on the text of a proposed rule.

HOW IS NEGOTIATED RULEMAKING DIFFERENT?

Generally a federal agency's staff drafts the text of a proposed rule. After circulation and comment within the agency, the rule will be printed in the Federal Register as a proposed rule. The public is then invited to comment on the rule. After public comment the agency may revise the rule to incorporate suggestions or eliminate problems identified as a result of public comment. The rule is then published in final form in the Federal Register and becomes effective on the date listed in the notice.

In a negotiated rulemaking proceeding, a well balanced group representing the regulated public, public interest groups, state and local governments, joins with a representative of the federal agency in a federally chartered advisory committee to negotiate the text of a rule before it is published as a proposed rule in the Federal Register. If the committee reaches consensus on the rule then the federal agency can use this consensus as a basis for its proposed rule. The proposed rule is still subject to public comment.

WHAT ARE THE ADVANTAGES OF NEGOTIATED RULEMAKING?

Federal agencies that have used negotiated rulemaking have identified several advantages to developing a rule by negotiation before notice and comment. The regulatory negotiation process allows the interested, affected parties a more direct input into the drafting of the regulation, thus ensuring that the rule is more sensitive to the needs and restrictions of both the parties and the agency. Rules drafted by negotiation have been found to be more pragmatic and more easily implemented at an earlier date, thus providing the public with the benefits of the rule while minimizing the negative impact of a poorly conceived or drafted regulation.

Because the negotiating committee has representatives of the major groups affected by or interested in the rule, the number of public comments is reduced. The tenor of public comment is more moderate. Fewer substantive changes are required before the rule is made final.

The committee can draw on the diverse experience and creative skills of the members to address problems encountered in writing a regulation. Many times the group together can propose solutions to difficult problems that no one member could have thought of or believed would work.

HOW DOES THE PROCESS WORK?

The federal agency establishes a formal advisory committee under the Federal Advisory Committee Act. A balanced mix of people is invited by the agency to participate and represent some identified interest or set of interests. Generally committees are composed of between 12 and 25 members representing both the public and private sectors. A neutral facilitator or mediator is used to convene the committee and to manage its meetings.

Meetings are announced in the Federal Register and are open to observation by members of the public. The number of meetings held depends on how complicated the rule is to draft, how much controversy there is amongst the committee members, and what the deadline is for the rule to be published and implemented.

Generally only the committee members speak during the meetings, although provisions are made for input by members of the audience. Caucuses can be called by committee members to speak with their constituency or with other members of the committee.

Decisions are made by consensus, not by majority vote. Consensus is defined by the committee prior to the start of its deliberations, however it is generally defined as an agreement by all parties that they can live with the provisions of the rule.

If consensus is reached, the agency will use it as a basis for their proposed rule. Committee members agree to support the rule as proposed if there are no substantive changes from the consensus agreement.

FOR ADDITIONAL INFORMATION ON REGULATORY NEGOTIATION AT EPA:

Chris Kirtz, Director, Consensus & Dispute Resolution Program, (202) 260-7565

Deborah Dalton, Deputy Director, Consensus & Dispute Resolution Program, (202) 260-5495

Briefing Paper for the Forum on Science in the National Interest
January, 1994

Alan G. Kraut, Ph.D.
Executive Director
American Psychological Society
1010 Vermont Ave. NW, # 1100
Washington, DC 20005-4907
202-783-2077 V 202-783-2083 F
bitnet akraut@aps

Priority Setting for the Federal Investment in Fundamental Science: *The Human Capital Initiative*

There is increasing criticism from Congress and even from some federal research agencies that the scientific community does not set priorities -- that the standard refrain from scientists is "just give us the money, we know what to do with it." The public also is more skeptical about science. We need to respond to our critics.

If we are honest, we will face the fact that a certain hubris has prevailed in science, with the result that we have squandered too much of the public's good will. That may have been more acceptable when the nation was pursuing single-minded goals and was freer spending. But we are in a different era now, one in which science must be mobilized on many fronts to address health and social problems, and one in which the context is a tightening federal fiscal belt with increasing pressure to justify how public funds are spent. In practical terms, we also need to recognize that our research monies are in competition with other legitimate and good programs -- low income housing, health benefits for the poor, veteran's programs, and Head Start, to name a few -- and priorities are being set *de facto* as Congressional appropriators and federal administrators make spending decisions. It is no longer a "given" that science will be funded for science's sake. Scientists must find ways to have systematic and ongoing input into the decisions that affect our research.

What follows is an overview of the effort of the behavioral science community to deal with these issues in developing the *Human Capital Initiative*, a national behavioral science research agenda that ties behavioral research to issues of national concern. This is offered in the hope that other areas of science can benefit from learning about how we approached the challenge of setting priorities and developing a common research agenda. Copies of the documents referred to in here are available at the above address.

The *Human Capital Initiative* (HCI) is the title of a document that was developed by representatives of some 70 behavioral science organizations over the course of two years. But the HCI also is an ongoing process. The HCI document targets six critical problems that face the nation today: productivity in the workplace; schooling and literacy; our aging society; drug and alcohol abuse; health and mental health; and violence. We make the case that each of these involves, at base, problems of human behavior, and each involves answering questions that require new "human capital" research -- basic and applied -- in behavioral science. The primary audience for the HCI is not the behavioral science community. It is deliberately written as a guide to our colleagues in federal research agencies, members of Congress, and decision makers in the private sector. But the ideas presented in the HCI are intended to appeal to researchers across a wide spectrum of

psychological specialties, and a key factor in the success of the overall process is that so many separate scientific organizations have come to agreement.

The HCI was written as an "umbrella" document, a framework for a research agenda. We currently are well into the second phase of HCI, in which specific research initiatives are being developed. These present what we know about the problem at hand, identify the critical research questions that remain and, using our best current judgement, identify the most promising areas for additional research on the problem. These initiatives are relatively brief, with a minimum of technical detail and jargon, and are intended to have a limited "shelf life," of a decade at most.

To date, two specific initiatives have been completed -- "The Changing Nature of Work" and "Vitality for Life: Psychological Research for Productive Aging." Others are under way in the areas of drug abuse, health and behavior, and mental illness. The mechanics of developing the initiatives are straightforward. There is a Human Capital Initiative Committee comprised of leading experts from various specialties within the field. This Committee selects a small number of individuals to develop an initial document and to serve as leaders for a drafting workshop on the chosen topic. A "Call for Participation" goes out to all interested organizations, inviting them to send a representative to the workshop. Participation is inclusive, not exclusive. Typically, the workshops are two days long, and have been attended by 30-40 representatives of different organizations who spend the time sitting together and hammering out the details of the specific initiative. A draft document is produced and circulated to the wider group of organizational representatives for their review and additional input. Following this review, the document is made final and distributed. Any participating organization may publish the document.

The HCI process continues to evolve, but already has had impact. Within behavioral science, the community has never been so united about its future, and ripple effects are being felt throughout the agencies that fund behavioral science. The National Institutes of Mental Health, Aging, Child Health and Human Development, Drug Abuse, and the National Science Foundation have underwritten the development of the HCI, and agency officials have taken part in all of the workshops. The impact also is being seen in Congress which has expressed appreciation for this approach. The HCI has received recognition several times from the Senate Appropriations Committee, which asked the National Science Foundation to use the HCI in planning its behavioral research programs. Similarly, the Senate Aging Committee asked the National Institute on Aging how it intends to incorporate the HCI in its planning. The National Institute of Mental Health is about to finish its own priority setting in basic behavioral science. The National Science Foundation is just starting a similar exercise, and is even using the title "Human Capital Initiative." And the National Institute on Drug Abuse approached us to help them set their priorities.

The complexity of the problems facing the nation requires new approaches. Research alone cannot solve the country's health and social problems, but neither will they be solved without systematic inquiry and painstaking analysis. A compelling case can be made for increased investment in research and it is the responsibility of the scientific community to make that case. The HCI is allowing the behavioral science community to do just that.